

Information content of weak lensing bispectrum: including the non-Gaussian error covariance matrix

Issha Kayo^{1*}, Masahiro Takada^{2†} and Bhuvnesh Jain^{3‡}

¹ *Department of Physics, Toho University, 2-2-1 Miyama, Funabashi, Chiba 274-8510, Japan*

² *Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), The University of Tokyo, Chiba 277-8582, Japan*

³ *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*

27 July 2012

ABSTRACT

We address a long-standing problem, how can we extract information in the non-Gaussian regime of weak lensing surveys, by accurate modeling of all relevant covariances between the power spectra and bispectra. We use 1000 ray-tracing simulation realizations for a Λ -CDM model and an analytical halo model. We develop a formalism to describe the covariance matrices of power spectra and bispectra of all triangle configurations, which extend to 6-point correlation functions. We include a new contribution arising from coupling of the lensing Fourier modes with large-scale mass fluctuations on scales comparable with the survey region via halo bias theory, which we call the halo sample variance (HSV) effect. We show that the model predictions are in excellent agreement with the simulation results for the power spectrum and bispectrum covariances. The HSV effect gives a dominant contribution to the covariance matrices at multipoles $l \gtrsim 10^3$, which arise from massive halos with masses $\gtrsim 10^{14} M_\odot$ and at relatively low redshifts $z \lesssim 0.4$. Since such halos are easy to identify from a multi-colour imaging survey, the effect can be estimated from the data. The bispectra add information to the power spectrum, and increase the total information content or the cumulative signal-to-noise ratio up to a maximum multipole of a few 10^3 , $(S/N)_{l_{\max}}$, by 20 – 50% compared to that of the power spectrum alone for upcoming surveys, which is equivalent to a factor of 1.4 – 2.3 larger survey area for the power spectrum measurement alone. However, the total information content of the power spectrum and bispectrum is still smaller than that for a Gaussian field by a factor of 3, mostly due to the HSV effect. Thus bispectrum measurements are useful for cosmology, but using information from upcoming surveys requires that non-Gaussian covariances are carefully estimated.

Key words: cosmology: theory — gravitational lensing — large-scale structure of the universe

1 INTRODUCTION

The accelerated expansion of the Universe is the most tantalizing problem in modern cosmology. Within Einstein’s gravity theory, General Relativity, cosmic acceleration can be explained by introducing dark energy, which acts as a repulsive, rather than attractive, force on the expansion of the universe. Alternatively, cosmic acceleration may be a signature of the breakdown of Einstein’s gravity theory on cosmological scales (Jain & Khoury 2010, for a review). Many on-going and upcoming wide-area galaxy surveys aim at testing dark energy and modified gravity scenarios as the origin of cosmic acceleration. In particular, ongoing and upcoming multi-colour imaging surveys include: the CFHT Weak Lensing Survey¹, the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS²), the VST Kilo-

* E-mail: kayo@ph.sci.toho-u.ac.jp

† E-mail: masahiro.takada@ipmu.jp

‡ E-mail: bjain@physics.upenn.edu

¹ <http://www.cfhtlens.org/astrophysicists/content-suitable-astrophysicists>

² <http://pan-starrs.ifa.hawaii.edu>

Degree Survey³, the Subaru Hyper Suprime-Cam Survey (Miyazaki et al. 2006)⁴, the Dark Energy Survey (DES)⁵, and in the next decade, the Large Synoptic Sky Survey (LSST)⁶, the ESA Euclid satellite mission⁷, and the NASA WFIRST satellite mission⁸.

Among different cosmological probes, weak gravitational lensing or cosmic shear is recognized as one of the most promising methods for constraining the nature of dark energy, provided systematic errors are well under control (see Bartelmann & Schneider 2001; Schneider 2006; Hoekstra & Jain 2008, for reviews). The bending of light rays emitted from a distant galaxy due to the foreground mass distribution causes the image to be distorted. The distortion signal is too weak for us to measure in single galaxies, but we can use a sufficiently large number of galaxy images, available from wide-field survey, to detect the correlated shear signals existing in-between different galaxy images. Weak lensing is a unique method of measuring the total matter distribution including dark matter, free of galaxy bias uncertainty, and allows a direct comparison of the measurement with theory that is in most case about the statistical properties of the dark matter distribution. The theoretical predictions are obtained using N-body simulations (e.g., Springel et al. 2006) and/or analytical approaches (e.g., Bernardeau et al. 2002; Cooray & Sheth 2002). The cosmological weak lensing signal has been measured by several groups (e.g., Hamana et al. 2003; Hoekstra & Jain 2008, and also see references therein), and we are waiting for measurements with much higher statistical precision from upcoming surveys.

Most previous works on weak lensing, in theoretical and observational studies, have focused on the shear two-point correlation function or equivalently its Fourier-transform, the power spectrum as the measurement method of weak lensing. The two-point correlation function contains the full information contained in the weak lensing field, *only if* the lensing field is Gaussian as in the cosmic microwave background (CMB) field (Komatsu et al. 2011). However this is not the case in reality. Nonlinear clustering in structure formation causes a coupling between different Fourier modes, and the mass density field at redshifts relevant for galaxy surveys is non-Gaussian. Since the lensing field is related to the mass density field via projection, the lensing field is also non-Gaussian. Hence, the power spectrum is not sufficient to extract the full information content of the lensing field. In fact, various studies have shown that the information content carried by the power spectrum might be saturated at multipole scales of a few 10^3 ; (Hamilton et al. 2006; Takahashi et al. 2009; Neyrinck et al. 2009, for the 3D mass density field) (Sato et al. 2009; Seo et al. 2011, for the lensing field), and (Lee & Pen 2008, for the result from the actual data). In particular, Sato et al. (2009) used 1000 ray-tracing simulation realizations for a Λ CDM model to study the power spectrum covariance and the information content of the power spectrum. They found that the information content is reduced by a factor of 2 at multipoles $l \simeq 10^3$ compared to the Gaussian case for a survey with typical source redshift $z_s \simeq 1$. Further, they showed that large-scale mass density fluctuations of scales outside the simulation area cause a significant contribution to non-Gaussian terms of the error covariance. They developed a formalism to describe the new non-Gaussian errors via the number fluctuations of massive halos based on halo bias theory, which we hereafter call the halo sample variance (HSV) effect (also see Hu & Kravtsov 2003; Takada & Bridle 2007).

Some fundamental questions remain unresolved: How important and useful are the non-Gaussian signals in the lensing field for cosmology? Which statistical method to extract the non-Gaussian signals is most useful? Can we recover the Gaussian information content, which should have existed in the linear field or the primordial field, by combining the power spectrum and the non-Gaussian signals? For weak lensing, there is additional expectation that the non-Gaussian signals will be useful for cosmology, because the skewness for example has been shown complementary to the power spectrum in its dependence on cosmological parameters (Bernardeau et al. 1997; Jain & Seljak 1997; Hui 1999; Jain et al. 2000; White & Hu 2000; Hamana & Mellier 2001; Van Waerbeke et al. 2001; Cooray & Hu 2001b; Takada & Jain 2002, 2004; Dodelson & Zhang 2005; Kilbinger & Schneider 2005; Semboloni et al. 2008; Bergé et al. 2010; Munshi et al. 2011; Pires et al. 2012). The attempt to measure the non-Gaussian signals from actual data was also made by several groups (Bernardeau et al. 2002; Zhang et al. 2003; Jarvis et al. 2004), although the detection was not significant yet.

In this paper, we study the lensing bispectrum – the Fourier-transformed counterpart of the three-point correlation function. It contains the lowest-order non-Gaussianity of the weak lensing field and is a natural extension of the two-point correlation function. We consider all the triangle configurations available from a given range of multipoles and include the full covariance matrix including the non-Gaussian error contributions up to the 6-point correlation functions. Gaussian errors were assumed in most previous work (Takada & Jain 2004; Martin et al. 2012). We use the 1000 simulation realizations to study the usefulness and complementarity of the lensing bispectrum compared to the power spectrum, and also develop an analytical formula to describe the bispectrum covariance for a given cosmology. In particular, we will show that the halo sample variance gives a significant contribution to the bispectrum covariance at $l \gtrsim$ a few 10^2 , and that the bispectrum does carry additional information to the power spectrum even in the presence of these significant correlations. Thus we will give a quantitative answer to the fundamental questions above.

This paper is organized as follows. After briefly reviewing the lensing power spectrum and bispectrum in § 2, we develop a formulation to describe the bispectrum covariance in § 3. In § 4, we show the main results: we study the bispectrum covariance using both simulations and analytical model predictions. We quantify the information content of the lensing bispectrum by including contributions from all the triangle configurations. § 5 is devoted to discussion and conclusion.

³ <http://www.astro-wise.org/projects/KIDS/>

⁴ <http://www.naoj.org/Projects/HSC/index.html>

⁵ <http://www.darkenergysurvey.org>

⁶ <http://www.lsst.org>

⁷ <http://sci.esa.int/science-e/www/area/index.cfm?fareaid=102>

⁸ <http://wfirst.gsfc.nasa.gov/>

2 PRELIMINARIES: LENSING POWER SPECTRUM AND BISPECTRUM

In the context of cosmological gravitational lensing, the convergence field is expressed as a weighted projection of the three-dimensional density fluctuation field between source and observer (see Bartelmann & Schneider 2001; Schneider 2006, for a thorough review)

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_H} d\chi W_{\text{GL}}(\chi) \delta[\chi, \chi\boldsymbol{\theta}], \quad (1)$$

where $\boldsymbol{\theta}$ is the angular position on the sky, χ is the comoving distance, and χ_H is the distance to the Hubble horizon. Note that we assume a flat geometry throughout this paper, and the radial distance χ is equivalent to the comoving angular diameter distance. The comoving distance $\chi(a)$ from an observer at $a = 1$ to a source at a is expressed in terms of the Hubble expansion rate $H(a)$ as $\chi(a) = \int_a^1 da' / [H(a')a'^2]$. For a single redshift of source galaxies, the lensing efficiency function $W_{\text{GL}}(\chi)$ is defined as

$$W_{\text{GL}}(\chi) \equiv \frac{3}{2} H_0^2 \Omega_{m0} a^{-1} \chi \left(1 - \frac{\chi}{\chi_s} \right), \quad (2)$$

where χ_s is the distance to the source galaxies. When including tomographic binning of source redshifts, see Eq. (4) in Takada & Jain (2009) for the lensing efficiency functions for different redshift bins. Under the flat-sky approximation, the Fourier transform of the lensing field is defined as

$$\kappa(\boldsymbol{\theta}) = \int \frac{d^2 \mathbf{l}}{(2\pi)^2} \tilde{\kappa}_{\mathbf{l}} e^{i\mathbf{l} \cdot \boldsymbol{\theta}}. \quad (3)$$

For a finite sky coverage of a survey, we need to use the discrete Fourier decomposition, rather than the infinite-range Fourier decomposition (also see Appendix in Takada & Bridle 2007, for the details).

In this paper, we study the bispectrum of lensing field and the covariance. The bispectrum is the Fourier-transformed counterpart of the three-point correlation function. The n -point power spectra relevant for the bispectrum covariance are defined in terms of the ensemble averages of the convergence fields in Fourier space as

$$\langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \rangle \equiv (2\pi)^2 P(\mathbf{l}_1) \delta_D^2(\mathbf{l}_1 + \mathbf{l}_2), \quad (4)$$

$$\langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \tilde{\kappa}_{\mathbf{l}_3} \rangle \equiv (2\pi)^2 B(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3) \delta_D^2(\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3), \quad (5)$$

$$\langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \tilde{\kappa}_{\mathbf{l}_3} \tilde{\kappa}_{\mathbf{l}_4} \rangle_c \equiv (2\pi)^2 T(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4) \delta_D^2(\mathbf{l}_1 + \mathbf{l}_2 + \mathbf{l}_3), \quad (6)$$

$$\langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \dots \tilde{\kappa}_{\mathbf{l}_n} \rangle_c \equiv (2\pi)^2 P_n(\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_n) \delta_D^2(\mathbf{l}_1 + \mathbf{l}_2 + \dots + \mathbf{l}_n) \quad \text{if } n \geq 5, \quad (7)$$

where $\delta_D^2(\mathbf{l})$ is the Dirac delta function; P , B and T are the lensing power spectrum, bispectrum and trispectrum, respectively; P_n is the n -point power spectrum. For the bispectrum covariance, we need to include up to the 6-point power spectra P_6 . Note that the higher-order correlation function than the bispectrum is the connected part of the n -point function, which characterize non-Gaussianity of the lensing field and cannot be expressed in terms of products of the power spectrum (and the lower-order correlation functions). Exchange symmetry of wavevectors, $\mathbf{l}_i \leftrightarrow \mathbf{l}_j$, reflects that the n -point correlation functions are invariant under permutations of the arguments. In addition, the delta function enforces translation invariance; adding constant vector into each argument vector \mathbf{l}_i does not change the n -point power spectrum. In other words, the vectors of $\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_n$ form the closed n -point polygonal-shape in Fourier space, i.e. $\mathbf{l}_1 + \mathbf{l}_2 + \dots + \mathbf{l}_n = \mathbf{0}$.

The lensing power spectrum and bispectrum can be given as the weighted line-of-sight projection of the three-dimensional power spectrum and bispectrum of the underlying mass distribution. Employing the Limber's approximation (Limber 1954) and the flat-sky approximation, the lensing power spectrum and bispectrum are expressed in terms of the three-dimensional mass spectra as

$$P(l) = \int_0^{\chi_H} d\chi W_{\text{GL}}^2(\chi) \chi^{-2} P_\delta \left(k = \frac{l}{\chi}; \chi \right), \quad (8)$$

$$B(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3) = \int_0^{\chi_H} d\chi W_{\text{GL}}^3(\chi) \chi^{-4} B_\delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \chi) |_{\mathbf{k}_i = \mathbf{l}_i / \chi}, \quad (9)$$

where P_δ and B_δ are the power spectrum and bispectrum of the matter distribution at each redshift $\chi (= \chi(z))$. Thus once the n -point spectra of the matter field are given for a given cosmological model, we can compute the n -point spectra of the lensing field. The above equation also means that statistical properties of the lensing spectra arise from those of the mass density field, since the prefactor functions such as the efficiency function $W_{\text{GL}}(\chi)$ is a pure geometrical factor, and not a statistical variable.

In reality, the power spectrum measurement is affected by the intrinsic shape noise due to a finite sampling of source galaxy shapes:

$$P^{\text{obs}}(l) = P(l) + \frac{\sigma_\epsilon^2}{\bar{n}_g}, \quad (10)$$

where σ_ϵ is the rms of intrinsic ellipticities per component, and \bar{n}_g is the mean number density of source galaxies per unit steradian. In the following, we will often omit the notation $^{\text{obs}}$ to refer $P^{\text{obs}}(l)$ for notational simplicity. The bispectrum is a measure of non-Gaussianity in the lensing field, so is not affected by the *random* shape noise.

3 LENSING COVARIANCE

The covariances of the lensing spectrum and bispectrum describe measurement accuracies of the spectra for a given survey. There are several sources of the measurement errors: the shot noise arising due to a finite sampling of galaxy shapes and the sample variance arising due to a finite survey area. If the lensing field is Gaussian, the different Fourier modes with $\mathbf{l} \neq \mathbf{l}'$ are independent, and therefore the sample variance is determined by the number of independent Fourier modes for a given multipole bin l that are available from the survey, yielding a simple formula of the sample variance contribution (e.g. Knox 1995). However, this is not case for the lensing field, because the lensing field is highly non-Gaussian at scales of interest (Takada & Jain 2004, 2009; Sato et al. 2009), and the different Fourier modes are correlated with each other. In the following, we discuss theory for the lensing covariance matrices.

3.1 Power spectrum covariance

The power spectrum covariance has been well studied by previous works (Scoccimarro et al. 1999; Cooray & Hu 2001a; Takada & Bridle 2007; Takada & Jain 2009; Sato et al. 2009). In particular, Sato et al. (2009) derived an expression of the power spectrum covariance including a new contribution from the large-scale mass density fields of scales outside survey area, and showed that the formula can well reproduce their ray-tracing simulation results. According to this work, the power spectrum covariance is given as

$$\text{Cov}[P(l_i), P(l_j)] = \frac{2\delta_{l_i l_j}^K}{N_{\text{pairs}}(l)} \left[P(l_i) + \frac{\sigma_\epsilon^2}{\bar{n}_g} \right]^2 + \frac{1}{\Omega_s} \int_{|\mathbf{l}| \in l_i} \frac{d^2 \mathbf{l}}{A(l_i)} \int_{|\mathbf{l}'| \in l_j} \frac{d^2 \mathbf{l}'}{A(l_j)} T(\mathbf{l}, -\mathbf{l}, \mathbf{l}', -\mathbf{l}') + \text{Cov}_{\text{HSV}}^{PP}(l_i, l_j; \Omega_s), \quad (11)$$

where $\delta_{l_i l_j}^K$ is the Kronecker delta, $\delta_{l_i l_j}^K = 1$ if $l_i = l_j$ within the bin width and otherwise $\delta_{l_i l_j}^K = 0$; Ω_s is the survey area in units of steradian; $A(l_i)$ is the area of the above integration in Fourier space, given as $A_s(l_i) \equiv \int_{|\mathbf{l}| \in l_i} d^2 \mathbf{l}$, where the integration range is confined to the wavevectors satisfying the condition $l_i - \Delta l/2 \leq |\mathbf{l}| \leq l_i + \Delta l/2$ (Δl is the bin width around the i -th bin, l_i); the third term on the r.h.s., denoted as Cov_{HSV} , is the new contribution which we call the halo sample variance (HSV) contribution (see below). The quantity $N_{\text{pairs}}(l_i)$ is the number of independent pairs of two vectors \mathbf{l} and $-\mathbf{l}$ in Fourier space, where the vector \mathbf{l} has the length l_i to within the bin width and “independent” here means different pairs discriminated by the fundamental Fourier mode of a given survey, $l_f \simeq 2\pi/\Theta_s$ (Θ_s is the angular scale of the survey area). At the limit $l_i \gg l_f$, $A(l_i) \simeq 2\pi l_i \Delta l$ and the number of independent Fourier modes is given as

$$N_{\text{pairs}}(l_i) \simeq \frac{2\pi l_i \Delta l}{(2\pi/\Theta_s)^2} = \frac{\Omega_s l_i \Delta l}{2\pi} = 2f_{\text{sky}} l_i \Delta l, \quad (12)$$

where f_{sky} is the sky fraction defined as $f_{\text{sky}} \equiv \Omega_s/4\pi$. See Takada & Bridle (2007) for a pedagogical derivation of the power spectrum covariance based on the discrete Fourier decomposition formulation (except for the third term $\text{Cov}_{\text{HSV}}^{PP}$). In Eq. (11), we ignored effects of non-trivial survey geometry for simplicity.

The first and second terms on the r.h.s. of Eq. (11) are standard covariance terms studied in most previous works. The first term describes the Gaussian covariance term that vanishes when $l_i \neq l_j$, i.e. no correlation between different multipole bins. The second term gives a non-Gaussian term arising from the lensing trispectrum (4-point correlation function), which describes the mode coupling between different multipole bins. If the lensing field is Gaussian as in the CMB field, the covariance has only the Gaussian term (the first term). Both the two terms scale with survey area as $1/\Omega_s$; the amplitudes decrease with increasing the survey area. It should also be noted that the Gaussian term depends on the multipole bin width, while the non-Gaussian terms not; a larger bin width relatively reduces the Gaussian term contribution at the multipole bin.

The third term of Eq. (11) arises from the mode coupling of the Fourier mode of our interest, l_i , with large-scale modes of scales comparable with or even outside the survey region. Such large-scale modes cannot be seen by an observer, but affect the power spectrum estimation as follows. For example, if the entire survey region happens to be in the overdensity region in the universe, which can be caused by the large-scale mass density fluctuations, the number of massive halos become to be larger than the ensemble average according to the halo bias theory, and vice versa (Mo & White 1996; Sheth et al. 2001). Thus the number of massive halos found in a finite survey volume becomes correlated with the mass density fluctuations of scales comparable with or larger than the survey field (Hu & Kravtsov 2003). To be more explicit, the number fluctuations of halos in the mass M and in the redshift slice centred at z are given as

$$\delta N(M) = b(M) \frac{d^2 V}{dz d\Omega} \Omega_s \Delta z \frac{dn}{dM} \bar{\delta}_m(\Theta_s; z), \quad (13)$$

where $d^2 V/dz d\Omega$ is the comoving volume per unit redshift interval and per unit solid angle, $d^2 V/dz d\Omega = \chi^2$ for a flat universe; dn/dM is the ensemble-averaged mass function of halos in the mass range $[M, M + dM]$; $b(M)$ is the halo bias parameter; $\bar{\delta}_m(\Theta_s)$ is the mass density fluctuation averaged within the survey volume in the redshift slice which has area Ω_s and the redshift width Δz . The lensing power spectrum amplitudes at small angular scales are sensitive to the number of massive halos in the survey region, as predicted in the halo model picture (Takada & Jain 2003a,b). Hence the lensing power spectrum amplitude becomes correlated with the large-scale density fluctuations via the number fluctuations of massive halos. The ensemble average of the power spectrum is not affected by the large-scale mode due to the fact $\langle \bar{\delta}_m(\Theta_s) \rangle = 0$. However, the correlation adds statistical scatters to the power spectrum measurement and therefore contribute to the power spectrum covariance, which is the HSV. At the limit $l_i, l_j \gg 1$ we are interested in, the HSV contribution is given as

$$\text{Cov}_{\text{HSV}}^{PP}(l_i, l_j; \Omega_s) = \int_0^{\chi_s} d\chi \left(\frac{d^2 V}{d\Omega d\chi} \right)^2 \left[\int dM \frac{dn}{dM} b(M) |\tilde{\kappa}_{l_i}|^2 \right] \left[\int dM' \frac{dn}{dM'} b(M') |\tilde{\kappa}_{l_j}|^2 \right] \left[\int_0^\infty \frac{k dk}{2\pi} P_m^L(k) |\tilde{W}(k\chi\Theta_s)|^2 \right], \quad (14)$$

where $\tilde{W}(l)$ is the Fourier transform of the survey window function; P_m^L is the linear mass power spectrum; $\tilde{\kappa}_l(\chi)$ is the Fourier transform of the convergence field for which we assume a Navarro-Frenk-White (NFW) halo (Navarro et al. 1997) (see Eq. 28 in Oguri & Takada 2011 or Section 3.2 in Takada & Jain 2003b for the expression of $\tilde{\kappa}_l$). Note that, for notational simplicity, we omitted to denote the redshift dependence of dn/dM , $b(M)$ and $\tilde{\kappa}_l$. In this paper, we simply consider the window function given by $\tilde{W}(x) = 2J_1(x)/x$, which corresponds to a circle-shaped survey geometry whose radius is Θ_s . For the halo model ingredients (dn/dM , $b(M)$ and the NFW profile), we will throughout this paper employ the same models as used in Takada & Jain (2009). Roughly speaking, the HSV term can be expressed as $\text{Cov}_{\text{HSV}}^{PP} \sim P_{1h}(l_i)P_{1h}(l_j)\sigma_m^2(\Omega_s)$, where P_{1h} is the 1-halo term of the lensing power spectrum and $\sigma_m(\Omega_s)$ is the rms of the projected linear mass density fluctuations smoothed with the angular scale of the survey area. As implied from the above equation, the HSV affects the band powers of different multipoles in the same way, and does not change the shape of the power spectrum. This HSV term cannot be realized as long as the discrete Fourier decomposition is used for deriving the power spectrum covariance, because the large-scale modes outside the survey region cannot be described by Fourier modes confined inside the survey region.

Another important feature is the HSV contribution depends on the survey area via the integration of the linear mass power spectrum, $\int k dk P_m^L(k) |\tilde{W}(k\chi\Theta_s)|^2$ unlike the other terms that scale as $1/\Omega_s$. For a power-law linear power spectrum, $P_m^L(k) \propto k^n$, the HSV term is found to scale as $\text{Cov}_{\text{HSV}}^{PP} \propto 1/(\Omega_s)^{1+n/2}$. Hence, for $n < 0$, which is indeed the case at $k \geq k_{\text{eq}}$ (k_{eq} is the wavenumber of the matter-radiation equality), the HSV amplitude more slowly decreases with increasing the survey area than the other terms do. On the other hand, when $n > 0$ or $k \leq k_{\text{eq}}$, the HSV term more quickly decreases with increasing the survey area. Thus the HSV term depends on the survey area in a non-trivial way, and we will study the relative importance of the HSV term for different survey areas.

Sato et al. (2009) showed that the HSV term gives a dominant contribution to the power spectrum covariance at large multipoles, $l \gtrsim 10^3$, over the first and second terms in Eq. (11). They showed that, once the HSV contribution is included, the model prediction given by Eq. (11) gives remarkably good agreement with the ray-tracing simulations for different multipole bins and a wide range of source redshifts.

There are even other sources of non-Gaussian errors arising from a correlation of the lensing field in the weakly nonlinear regime with the mass density fluctuations of scales comparable with or larger than the survey area. This contribution can be formulated based on the perturbation theory of structure formation, which is valid for the mass density field in the weakly nonlinear regime. Rimes & Hamilton (2005) first studied this effect for the 3D mass power spectrum, and named this new contribution the beat-coupling mode (also see Hamilton et al. 2006; Sefusatti et al. 2006; Takahashi et al. 2009). Furthermore, de Putter et al. (2012) recently pointed out that the large-scale density fluctuations cause an apparent modulation in the mean density estimated from a finite survey region, and cause even additional negative contribution to the covariance. This term was shown to have a similar amplitude to the beat-coupling mode. These large-scale mode contribution, which is relevant for the weakly nonlinear regime, differs from the HSV effect, and Takada & Jain (2009) showed that the contribution to the lensing power spectrum covariance is negligible compared to the non-Gaussian errors arising from the 4-point function (the second term in Eq. 11) at multipoles of $l \gtrsim$ a few 10^2 . Hence, in this paper, we ignore the non-Gaussian errors arising from the mode-coupling in the weakly nonlinear regime for simplicity.

3.2 Bispectrum covariance

As we discussed above, the lensing bispectrum is given as a function of triangle configurations. An estimator of the lensing bispectrum from the finite-area lensing survey can be found, by extending the method developed in Takada & Bridle (2007) (see Appendix A for the details), as

$$\hat{B}(l_1, l_2, l_3) = \frac{1}{\Omega_s N_{\text{trip}}(l_1, l_2, l_3)} \sum_{\mathbf{q}_i} \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3), \quad (15)$$

where $\mathbf{q}_{123} \equiv \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$ and the summation runs over all the pixels of \mathbf{q}_1 , \mathbf{q}_2 and \mathbf{q}_3 . The function $\Delta_{\mathbf{q}_{123}}$ denotes the selection function defined so that it gives unity if each vector has a target length such as $l_i - \Delta l_i/2 \leq q_i \leq l_i + \Delta l_i/2$ ($i = 1, 2, 3$) and the three vectors form the triangle configuration in Fourier space, $\mathbf{q}_{123} = \mathbf{0}$; otherwise $\Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) = 0$. The quantity N_{trip} is the number of the triplets in Fourier space that form a given triangle configuration specified by three side lengths l_1, l_2, l_3 within the bin widths. This is calculated in terms of the selection function as $N_{\text{trip}} = \sum_{\mathbf{q}_i} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3)$. The prefactor $1/\Omega_s$ is from our definition of the discrete Fourier decomposition (see Takada & Bridle 2007).

As derived in Joachimi et al. (2009) and we derive in Appendix A, for the limit of large multipole bins, $l_1, l_2, l_3 \gg l_f$, we can analytically estimate the number of independent combinations of triplets ($\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$) satisfying a given triangle configuration (where we meant by “independent” the triplets are discriminated by the fundamental mode of a given survey area):

$$N_{\text{trip}}(l_1, l_2, l_3) \equiv \sum_{\mathbf{q}_i: q_i \in l_i} \Delta_{\mathbf{q}_{123}} \simeq \frac{\Omega_s^2 l_1 l_2 l_3 \Delta l_1 \Delta l_2 \Delta l_3}{2\pi^3 \sqrt{2l_1^2 l_2^2 + 2l_1^2 l_3^2 + 2l_2^2 l_3^2 - l_1^4 - l_2^4 - l_3^4}}, \quad (16)$$

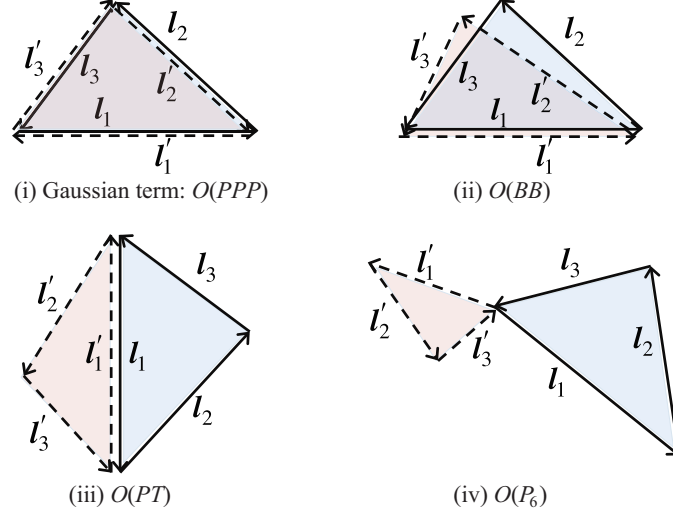


Figure 1. Illustration of the different terms of the bispectrum covariance. The triangle configurations for two bispectra in the covariance matrix, $B(l_1, l_2, l_3)$ and $B(l'_1, l'_2, l'_3)$, are specified by sets of the three vectors (l_1, l_2, l_3) or (l'_1, l'_2, l'_3) , denoted by the solid and dashed lines, respectively. The three vectors satisfy the triangle conditions $l_1 + l_2 + l_3 = \mathbf{0}$ and $l'_1 + l'_2 + l'_3 = \mathbf{0}$. (i) The Gaussian part of the bispectrum covariance, which arises only if the two triangle configurations have the same shape (within the coarseness of the bin widths). Hence the Gaussian amplitude is proportional to $P(l_1)P(l_2)P(l_3)$ and contributes to the diagonal terms of the covariance matrix. The two vectors with the same length such as l_1 and l'_1 and in the opposite directions yield the power spectrum after the ensemble average. (ii) One non-Gaussian part of the bispectrum covariance which arises if one side of the two triangles is in the same length and in the same direction; here $l_1 = l'_1$ is shown as an example. The amplitude of this covariance term is proportional to $B(l_1, l_2, l_3)B(l'_1, l'_2, l'_3)$. (iii) A second non-Gaussian part that arises if one side length of the two triangles is the same but in the opposite direction, $l_1 = -l'_1$. The amplitude is $O(PT)$; in particular, it is given by $P(l_1)T(l_1, l_3, l'_2, l'_3)$ in this case. The surrounding four vectors (l_2, l_3, l'_2, l'_3) in this case) form a quadrangular configuration satisfying the condition $l_2 + l_3 + l'_2 + l'_3 = \mathbf{0}$, which gives the trispectrum contribution. (iv) A third non-Gaussian part which arises for generic triangle configurations and therefore contributes to all diagonal and off-diagonal terms of the covariance matrix. The amplitude is proportional to the connected part of the 6-point correlation function, P_6 . As indicated, the 6 vectors of $(l_1, l_2, l_3, l'_1, l'_2, l'_3)$ arise from the two triangles that form the 6-point configuration in Fourier space, although the two vertices of the 6 points are collapsed to one point due to the triangle conditions.

where Δl_i is the bin width of the i -th side length. For the small-angle scales $l_i \gg 1$ (flat-sky approximation limit) we are interested in, the equation above gives a good approximation to the Wigner-3j symbols that appears in the bispectrum covariance derived under the full-sky approach (see around Eq. 16 in Takada & Jain (2004)).

The bispectrum covariance can be similarly defined as

$$\begin{aligned} \text{Cov}[B(l_1, l_2, l_3), B(l'_1, l'_2, l'_3)] &\equiv \langle \hat{B}(l_1, l_2, l_3) \hat{B}(l'_1, l'_2, l'_3) \rangle - B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3) \\ &= \frac{1}{\Omega_s^2 N_{\text{trip}}(l_1, l_2, l_3) N_{\text{trip}}(l'_1, l'_2, l'_3)} \sum_{\mathbf{q}_i; q_i \in l_i} \sum_{\mathbf{q}'_i; q'_i \in l'_i} \left[\langle \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \tilde{\kappa}_{\mathbf{q}'_1} \tilde{\kappa}_{\mathbf{q}'_2} \tilde{\kappa}_{\mathbf{q}'_3} \rangle \Delta \mathbf{q}_{123} \Delta \mathbf{q}'_{123} \right] - B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3). \end{aligned} \quad (17)$$

Thus the bispectrum covariance arises from the 6-point correlation function of the lensing field, and is given as a function of the two different triangle configurations for the two bispectrum, $B(l_i)$ and $B(l'_i)$, respectively.

We present the detailed derivation of the bispectrum covariance in Appendix A based on the discrete Fourier decomposition formulation. Here we just give the expression of the bispectrum covariance, which has three contributions; the Gaussian and non-Gaussian errors and the HSV contribution, respectively:

$$\begin{aligned} \text{Cov}[B(l_1, l_2, l_3), B(l'_1, l'_2, l'_3)] &= \text{Cov}_{\text{Gauss}} + \text{Cov}_{\text{NG}} + \text{Cov}_{\text{HSV}}^{BB} \\ &= \frac{\Omega_s}{N_{\text{trip}}(l_1, l_2, l_3)} P(l_1) P(l_2) P(l_3) \left[\delta_{l_1 l'_1}^K \delta_{l_2 l'_2}^K \delta_{l_3 l'_3}^K + \delta_{l_1 l'_1}^K \delta_{l_2 l'_3}^K \delta_{l_3 l'_2}^K + \delta_{l_1 l'_2}^K \delta_{l_2 l'_1}^K \delta_{l_3 l'_3}^K + 3 \text{ perms.} \right] \\ &\quad + \frac{2\pi}{\Omega_s} B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3) \left[\frac{\delta_{l_1 l'_1}^K}{l_1 \Delta l_1} + \frac{\delta_{l_1 l'_2}^K}{l_1 \Delta l_1} + 7 \text{ perms.} \right] \\ &\quad + \delta_{l_1 l'_1}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_2, l'_3) + \delta_{l_1 l'_2}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_1, l'_3) + 7 \text{ perms.} \\ &\quad + \frac{1}{\Omega_s} \int \frac{d\psi}{2\pi} P_6(l_1, l_2, l_3, l'_1, l'_2, l'_3; \psi) \\ &\quad + \text{Cov}_{\text{HSV}}^{BB} \end{aligned} \quad (18)$$

where the notation “NG” stands for the “non-Gaussian” error contribution, and P_6 denotes the connected part of the 6-point correlation function. Fig. 1 shows a diagram picture of these covariance terms from the first to the fourth lines on the r.h.s. of Eq. (18). When further

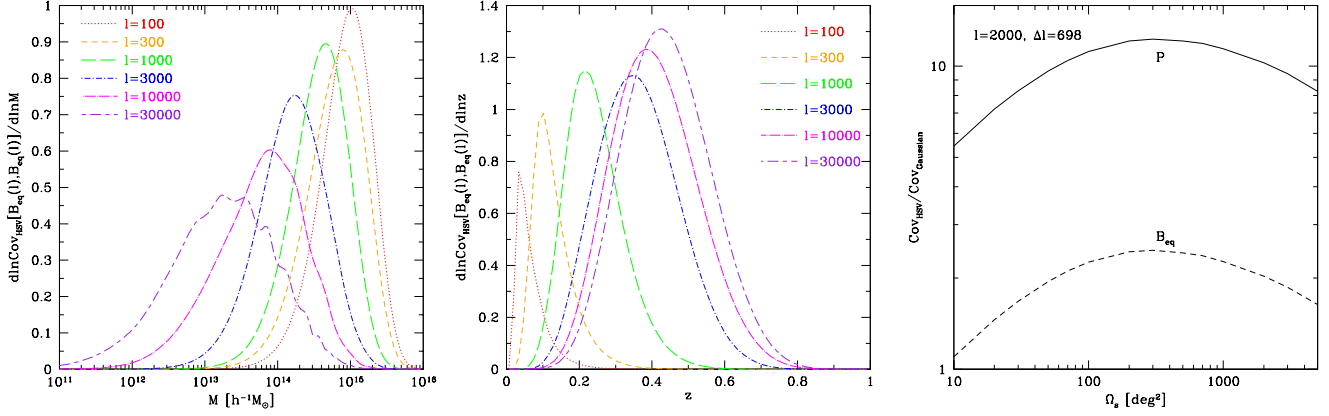


Figure 2. *Left panel:* The plot shows how halos of different masses contribute to the halo sample variance (HSV) terms of the bispectrum covariance for a given equilateral triangle configuration. We used the halo model expression (Eq. 19) to compute the fractional contributions. For the triangle with side length $l = 1000$, halos with masses $\gtrsim 10^{14} M_\odot$ dominate the contribution to the HSV. *Middle:* Similar plot, but for the redshift distribution of the bispectrum covariance. For the triangle with $l = 1000$, halos at $z \lesssim 0.4$ give the dominant contribution. *Right:* The survey area dependence of the HSV contribution for the power spectrum and the bispectrum, relative to that of the Gaussian error, which scales as $\text{Cov}_{\text{Gaussian}} \propto 1/\Omega_s$. Here we consider the scale of $l = 2000$ for $P(l)$ and $B_{\text{eq}}(l)$, employ the flat sky approximation, and assume the survey geometry given by $\Omega_s = \pi\Theta_s^2$.

including the intrinsic shape noise contribution, we just replace the power spectra in the terms of the above equation ($O(P^3)$ and $O(PT)$ terms) with the power spectrum including the shot noise contribution (Eq. 10).

The terms of the first line on the r.h.s. are the Gaussian covariance terms, which contribute only to the diagonal terms of the bispectrum covariances. The combination of the Kronecker deltas $\delta_{l_1 l'_1}^K \delta_{l_2 l'_2}^K \delta_{l_3 l'_3}^K$ is non-vanishing only if the two triangle configurations are in the same “shape” within the bin widths. In particular, if triangle configurations have symmetry such as isosceles or equilateral triangles, the combination of Kronecker deltas (the terms in the square bracket on the first line) yield a factor of 2 or 6 for isosceles and equilateral triangles, respectively. The factors account for the fact that different triangles transformed by parity and permutation transformations ($l_i \leftrightarrow l'_j$) are not independent for a statistically homogeneous and isotropic field. For a general triangle configuration $l_i \neq l'_j$, the factor becomes unity. The prefactor $N_{\text{trip}}(l_1, l_2, l_3)$ in the first term is given by Eq. (16).

The terms in the second- and to the fourth lines on the r.h.s. of Eq. (18) are the non-Gaussian error contributions, which arise from the higher-order correlation functions of the lensing field. The coefficient of each term such as $2\pi/(\Omega_s l_1 \Delta l_1)$ is from the number of independent configurations in Fourier space that form a given configuration of 6 wavevectors $(l_1, l_2, l_3, l'_1, l'_2, l'_3)$ in Fig. 1 (also see Appendix A for the mathematical derivation). The angular integration in the term including P_6 is over the angle ψ between the vectors l_1 and l'_1 in order to include contributions over all the possible 6-point configurations in Fourier space in Fig. 1. To be more precise, for a given element of the bispectrum covariance matrix, each of the two triangle configurations for the two bispectra is fixed, and the 6-point configuration can be parametrized by the angle ψ . Note that all the terms in the first, second and third lines on the r.h.s. depend on the multipole bin widths such as Δl_1 , while the term including P_6 and the HSV do not depend on the bin widths.

As in the power spectrum covariance (Eq. 11), the number fluctuations of massive halos due to the large-scale mass fluctuations affect the bispectrum estimated from a finite area survey. Similarly we can find that the HSV contribution to the bispectrum covariance to be given as

$$\begin{aligned} \text{Cov}[B_\kappa(l_1, l_2, l_3), B_\kappa(l'_1, l'_2, l'_3)]_{\text{HSV}} &= \int d\chi \left(\frac{d^2 V}{d\chi d\Omega} \right)^2 \left[\int dM \frac{dn}{dM} b(M) \tilde{\kappa}_M(l_1) \tilde{\kappa}_M(l_2) \tilde{\kappa}_M(l_3) \right] \\ &\times \left[\int dM' \frac{dn}{dM'} b(M') \tilde{\kappa}_{M'}(l'_1) \tilde{\kappa}_{M'}(l'_2) \tilde{\kappa}_{M'}(l'_3) \right] \int \frac{k dk}{2\pi} P_m^L(k; \chi) |\tilde{W}(k\chi\Theta_s)|^2. \end{aligned} \quad (19)$$

The HSV contribution to the bispectrum covariance, more generally the higher-order covariance, has not been realized in previous works. By comparing with ray-tracing simulations, we will show below that the HSV contribution is dominant over other covariance terms at $l \gtrsim 1000$, in the nonlinear regime, and adding the HSV term to the model predictions significantly improves agreement with the simulation results. Again notice that the HSV term scales with the survey area in a non-trivial way via the linear mass power spectrum (see the discussion below Eq. 14), while the other terms scale with survey area as $1/\Omega_s$.

The left and middle panels of Fig. 2 show which halos of mass and redshift range contribute to the HSV term of the bispectrum covariance for an equilateral triangle configuration of a given side length. For the triangle with $l = 1000$, halos with masses $\gtrsim 10^{14} M_\odot$ and at redshift $z \lesssim 0.4$ give a dominant contribution to the HSV effect. These halos are relatively easy to identify from a concentration of galaxies on the sky, X-ray or the Sunyaev-Zel’dovich effect. In other word, identifying such massive halos in a survey region and comparing the number with the expected number for a fiducial cosmological model will help understand the HSV effect for the weak lensing observables in the survey region (also see Takada & Bridle 2007, for the similar discussion). The right panel shows how the HSV terms scale with the

survey area, in comparison with that of the Gaussian covariance term for the power spectrum covariance and the bispectrum covariance for equilateral triangles. Here we considered the particular multipole bin $l = 2000$; $P(l)$ and $B_{\text{eq}}(l)$, and assumed the flat-sky approximation and the survey geometry $\Omega_s = \pi\Theta_s^2$ for simplicity. If the curves are constant or independent of Ω_s , it means that the HSV term scales as $\text{Cov}_{\text{HSV}} \propto 1/\Omega_s$ as in the other covariance terms. Hence the plot shows that, with increasing Ω_s the HSV term more slowly decreases for $\Omega_s \lesssim$ a few 100 deg^2 or the larger survey area than the other terms do.

3.3 Cross-covariance between power spectrum and bispectrum

Since the power spectrum and bispectrum are not independent for the non-Gaussian field, we need to account for the cross-covariance between the two observables in order not to double-count the information content.

Similarly, we can derive the cross-covariance (see Appendix B):

$$\begin{aligned} \text{Cov} [P^{\text{est}}(l), B^{\text{est}}(l_1, l_2, l_3)] &= \text{Cov}_{NG}^{PB} + \text{Cov}_{\text{HSV}}^{PB} \\ &= \delta_{ll_1}^K \frac{4\pi}{\Omega_s l_1 \Delta l_1} P(l) B(l, l_2, l_3) + 2 \text{ perms.} + \frac{1}{\Omega_s} \int \frac{d\psi}{2\pi} P_5(\mathbf{l}, -\mathbf{l}, \mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3; \psi) + \text{Cov}_{\text{HSV}}^{PB}, \end{aligned} \quad (20)$$

where P_5 is the 5-point correlation function, and ψ is the angle between the vectors \mathbf{l} and \mathbf{l}_1 as in Eq. (18). The HSV term is similarly given as

$$\begin{aligned} \text{Cov}[P_\kappa(l), B_\kappa(l_1, l_2, l_3)]_{\text{HSV}} &= \int d\chi \left(\frac{d^2 V}{d\chi d\Omega} \right)^2 \left[\int dm \frac{dn}{dM} b(M) |\tilde{\kappa}_M(l)|^2 \right] \\ &\quad \times \left[\int dM \frac{dn}{dM'} b(M') \tilde{\kappa}_{M'}(l_1) \tilde{\kappa}_{M'}(l_2) \tilde{\kappa}_{M'}(l_3) \right] \int \frac{kd k}{2\pi} P_m^L(k; \chi) |\tilde{W}(k\chi\Theta_s)|^2. \end{aligned} \quad (21)$$

We will use these equations when computing the total information content for a combined measurement of the lensing power spectrum and bispectrum for a given survey.

3.4 Halo sample variance contribution to the n -point correlation function measurement

As we have discussed, the HSV effect contributes to the lensing spectrum covariances (see Eqs. 14, 19 and 21). All the equations have similar forms at the small-angle limit, where the 1-halo term is dominated in the halo model picture. Extending these findings, we can find the HSV contribution to the covariance matrix between any n - and n' -point correlation functions in Fourier space:

$$\begin{aligned} \text{Cov}[P_n(l_1, l_2, \dots, l_n), P_{n'}(l'_1, l'_2, \dots, l'_{n'})]_{\text{HSV}} &= \int d\chi \left(\frac{d^2 V}{d\chi d\Omega} \right)^2 \left[\int dM \frac{dn}{dM} b(M) \tilde{\kappa}_M(l_1) \tilde{\kappa}_M(l_2) \dots \tilde{\kappa}_M(l_n) \right] \\ &\quad \times \left[\int dM' \frac{dn}{dM'} b(M') \tilde{\kappa}_{M'}(l'_1) \tilde{\kappa}_{M'}(l'_2) \dots \tilde{\kappa}_{M'}(l'_{n'}) \right] \int \frac{kd k}{2\pi} P_m^L(k; \chi) |\tilde{W}(k\chi\Theta_s)|^2. \end{aligned} \quad (22)$$

Roughly speaking, the amplitude of the HSV term simply scales as $\text{Cov}_{\text{HSV}} \sim P_n^{1h}(l_i) P_{n'}^{1h}(l'_j) \sigma_m^2(\Theta_s)$. Thus any n -point correlation functions at small angle scales can be affected by the large-scale mass fluctuations of scales comparable with or outside the survey area.

We should also emphasize that the HSV contribution affects a measurement(s) of any two- or three-dimensional correlation functions from a finite area survey, and can be very important if one is interested in the small-scale signals which are sensitive to the abundance of halos in the finite survey region (e.g., see Shaw et al. 2009; Zhang & Sheth 2007, a similar discussion on the SZ power spectrum measurement).

3.5 Halo model predictions for the lensing covariances

As we have described up to the preceding section, the power spectrum and bispectrum covariance calculations require to compute the 4-, 5- and 6-point correlation functions in addition to the power spectrum and bispectrum. For the power spectrum and bispectrum, there are some efforts developing secure model predictions in combination with simulations, e.g., Smith et al. (2003); Valageas & Nishimichi (2011a) for the power spectrum and Scoccimarro & Frieman (1999); Valageas & Nishimichi (2011b) for the bispectrum. The higher-order functions are, however, fairly uncertain because there are fewer studies to compare the model predictions with simulations (e.g. Takada & Jain 2002, for an attempt to compute the kurtosis which is the collapsed 4-point function), partly because the higher-order correlations require a substantial amount of computational costs. Instead of pursuing a reliable model for the higher-order correlation functions, in this paper we employ the halo model approach to compute the higher-order functions (Seljak 2000; Peacock & Smith 2000; Ma & Fry 2000; Scoccimarro et al. 2001; Cooray & Sheth 2002) in which the correlations of the mass distribution are expressed as two separate contributions: correlation of dark matter particles within the same halo, and correlations between particles in different halos. We have found that, up to the 4-point correlation functions, the halo model predictions are accurate at 10 – 30% level in the amplitude compared to N-body simulations (in particular for lensing fields; Takada & Jain 2002, 2003a,b). Since our purpose of this paper is to assess the importance of non-Gaussian error contributions to the bispectrum covariance matrices, we consider that the halo model approach is adequate enough.

We know that most of the lensing information come from the small angle scales in the nonlinear clustering regime to which the 1-halo term, the correlation arising from the same halo, provides a dominant contribution. In addition, the non-Gaussian errors are important only at the small scales, as can be explicitly found from Fig. 5 in Sato et al. (2009). For these reasons, we include only the 1-halo terms to compute the non-Gaussian error contributions to the lensing covariances, which significantly simplifies the computation. Although the n -point correlation function depends on n wavevectors such as $P_n(l_1, l_2, \dots, l_n)$, the 1-halo term does not depend on any angle between the vectors, but rather depends only on the length of each vector; $P_n^{1h}(l_1, l_2, \dots, l_n) = P_n^{1h}(l_1, l_2, \dots, l_n)$, reflecting spherical mass distribution around halo in a statistical average sense. To be more explicit, assuming the Limber's approximation, the 1-halo term of the n -point correlation function can be computed as follows (see around Eq. 30 in Takada & Jain 2003a):

$$P_n^{1h}(l_1, l_2, \dots, l_n) = \int_0^{\chi_s} d\chi \frac{d^2 V}{d\chi d\Omega} \int dM \frac{dn}{dM} \tilde{\kappa}_M(l_1) \tilde{\kappa}_M(l_2) \cdots \tilde{\kappa}_M(l_n). \quad (23)$$

The 1-halo term can be computed by a two-dimensional numerical integration, as the Fourier transform of the lensing field due to an NFW halo, $\tilde{\kappa}_M(l)$, can be computed analytically for a given halo with mass M and at redshift z .

As can be found from Eqs. (11), (18) and (20), some of the non-Gaussian terms can be further simplified; e.g., one of the non-Gaussian terms in the power spectrum covariance (Eq. 11) can be simplified as

$$\frac{1}{\Omega_s} \int_{|l| \in l_i} \frac{d^2 l}{A(l_i)} \int_{|l'| \in l_j} \frac{d^2 l'}{A(l_j)} T(l, -l, l', -l') \simeq \frac{1}{\Omega_s} T^{1h}(l_i, l_i, l_j, l_j), \quad (24)$$

where we have assumed that the lensing trispectrum does not largely change within the multipole bin Δl around the bins l_i and l_j . Thus the above approach allows a faster computation of the power spectrum and bispectrum covariances.

4 RESULTS: COMPARISON WITH RAY-TRACING SIMULATIONS

4.1 Ray-tracing simulations

To study the lensing covariance matrices, we use 1000 realizations of ray-tracing simulations for a Λ CDM model in Sato et al. (2009). Although the simulations were done for various source redshifts ranging from $z_s = 0.6$ to 3, we use the outputs of $z_s = 1$ in this paper. In brief, each realization has area of 5×5 square degrees ($\Omega_s = 0.0076$ steradian) in square shaped geometry. The Λ CDM model adopted is characterized by cosmological parameters: the matter density parameter $\Omega_m = 0.238$, the baryon density parameter $\Omega_b = 0.042$, the initial spectral index $n_s = 0.958$, the amplitude of the density fluctuations $\sigma_8 = 0.76$ and the Hubble constant of $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.732$. The linear matter power spectrum used to set the initial conditions of N-body simulations is computed by the public code CAMB (Lewis et al. 2000). It was shown that the ray-tracing simulations are reliable to within a 5% accuracy up to multipole $l \simeq 6000$ for the power spectrum and up to $l \simeq 4000$ for the bispectrum. See Sato et al. (2009) and Valageas et al. (2012) for more details⁹.

As can be found from Fig. 1 in Sato et al. (2009), the ray-tracing simulations were done in a light cone of area 5×5 square degrees, viewed from an observer position ($z = 0$ position). The projected mass density fields in intermediate redshift slices were generated from N-body simulations whose size include the region bigger than the volume covered by the light cone. Hence the lensing fields have contributions from the mass density field of scales outside the ray-tracing simulation area, although, exactly speaking, the modes outside the N-body simulation box were not included. Thus the ray-tracing simulations allow us to study the HSV effect on the covariance matrices.

4.2 Measuring power spectrum, bispectrum and the covariance matrices from simulations

In each simulation realization, the lensing convergence field, $\kappa(\theta)$, is given on 2048×2048 grids. We used the FFT method to compute the Fourier-transformed field, $\tilde{\kappa}(l)$. The fundamental mode of the discrete Fourier decomposition is $l_f = 72 (= 2\pi/5^\circ)$ and the Nyquist frequency is ~ 70000 , which is large enough compared to the resolution limit of the N-body simulations. We use multipole bins that are logarithmically spaced by $\Delta \ln l = 0.3$ ($\Delta \log l \simeq 0.13$), which significantly reduces the number of triangle configurations (the number of different bispectra) we need to consider, compared to $\Delta l = 1$ as in the CMB case. We consider 16 multipole bins in total; the 1st bin is in the range $l = [72, 97.2]$, and the 16th bin in the range $l = [6481.2, 8748.8]$, so $l_{\min} = 72$ and $l_{\max} = 8748.8$.

The power spectrum of the i -th multipole bin l_i , $P(l_i)$, is estimated from each realization by computing an azimuthal average of the estimator, $\tilde{\kappa}_l \tilde{\kappa}_{-l}$, where $|l|$ resides in the target bin. We then averaged the estimated power spectra in 1000 realizations to estimate the ensemble-averaged power spectrum (corresponding to the power spectrum for the area of $1000 \times 25 = 25000$ square degrees). Then we computed the scatters among the power spectra of 1000 realizations in order to estimate the covariance matrix of the power spectra (see Sato et al. 2009, for details). Hence the covariance matrix is for an area of 25 square degrees.

The bispectrum is given as a function of triangle configuration. We use three side lengths (l_1, l_2, l_3) to parametrize triangle configuration, where the triangle conditions are given as $|l_j - l_k| \leq l_i \leq l_j + l_k$. Although the multipole bin used has a logarithmically-spaced bin width,

⁹ The simulation data is available at <http://www.a.phys.nagoya-u.ac.jp/~masanori/HSC/>

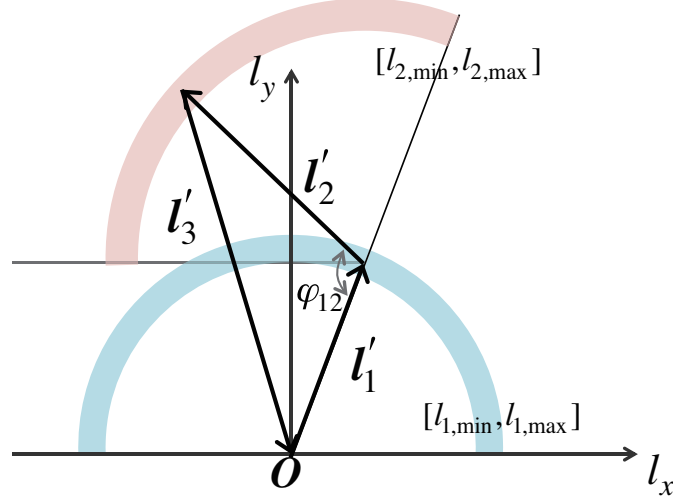


Figure 3. Illustration describing how to choose three vectors in Fourier space that satisfy triangle configurations within multipole bin widths. The triangle configuration is specified by three side lengths which have central values (l_1, l_2, l_3) and the widths $\Delta \ln l = 0.3$. The three vertices chosen are used in estimating the bispectrum from ray-tracing simulations by averaging the estimator $\text{Re}[\tilde{\kappa}_{\mathbf{l}'_1} \tilde{\kappa}_{\mathbf{l}'_2} \tilde{\kappa}_{\mathbf{l}'_3}]$ over the triangles (see text for details). First, the first vector \mathbf{l}'_1 is chosen from the annulus (shaded region), which is in the upper half plane and has the radius in the range $[l_{1,\min}, l_{1,\max}]$ (a bin of l_1). Then, the second vector \mathbf{l}'_2 is chosen from the annulus which has position angle in the range $\arg(\mathbf{l}'_1) \leq \arg(\mathbf{l}'_2) \leq \pi$ and has radius in the range $[l_{2,\min}, l_{2,\max}]$. For the given pair of \mathbf{l}'_1 and \mathbf{l}'_2 , the third vector \mathbf{l}'_3 is determined by the triangle condition $\mathbf{l}'_3 = -\mathbf{l}'_1 - \mathbf{l}'_2$; if \mathbf{l}'_3 is in the range $l_{\min} \leq l'_3 \leq l_{\max}$, the triplet of $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$ is accepted and otherwise discarded. The angle φ_{12} is used for discussion in Appendix A.

we impose the triangle conditions on the central values of the multipole bins. In addition we impose the condition $l_1 \leq l_2 \leq l_3$ so that every triangle configuration is counted once. For the 16 multipole bins above, we have 204 triangle configurations in total.

For a triangle configuration that is specified by the side lengths (l_1, l_2, l_3) (with the bin widths), we can estimate the bispectrum from each ray-tracing simulation by averaging the estimator $\text{Re}[\tilde{\kappa}_{\mathbf{l}'_1} \tilde{\kappa}_{\mathbf{l}'_2} \tilde{\kappa}_{\mathbf{l}'_3}]$ over all the triplets $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$ which satisfy the triangle conditions; the length of each vector is in the triangle bin such as $l_{\min} \leq l'_1 \leq l_{\max}$. Note that, as long as the triplets of $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$ in Fourier space have the same side lengths l_1, l_2, l_3 within the bin widths, all the triangles transformed by parallel translation, rotation and parity transform are equivalent to yield the same bispectrum due to rotation, parallel translation and parity invariance for a statistically homogeneous and isotropic field. In our simulations, the Fourier-transformed convergence field, $\tilde{\kappa}(\mathbf{l})$, is given on 2048×2048 grids in Fourier space, where the grids are linearly spaced by the fundamental mode, $l_f = 2\pi/\Theta_s = 72$. To have an efficient computation over 1000 realizations, we first built the table of three vectors (grids), $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$, where each triplet satisfies the triangle conditions ($|\mathbf{l}'_j - \mathbf{l}'_k| \leq l'_i \leq |\mathbf{l}'_j + \mathbf{l}'_k|$) and is assigned to one of the triangle configurations binned by three side lengths (l_1, l_2, l_3) . Then we used the *same* table of triplets for the 1000 realizations to compute the average and scatters of the estimated bispectra as a function of triangle configurations.

To be more precise, we built the table of three vectors (grids) in the way illustrated in Fig. 3. First, we choose the first vector \mathbf{l}'_1 from one of the grids in the upper half of Fourier space, i.e. $0 \leq \arg(\mathbf{l}'_1) < \pi$, by imposing the condition that the length l'_1 is in the range of the multipole bin, $l_{\min} \leq l'_1 < l_{\max}$. Then we survey for the second vector \mathbf{l}'_2 from the region where the length is $l_{\min} \leq l'_2 < l_{\max}$ and the position angle satisfies $\arg(\mathbf{l}'_1) \leq \arg(\mathbf{l}'_2) < \pi$ (more precisely, $l'_1 \leq l'_2$ if and only if $\arg(\mathbf{l}'_1) = \arg(\mathbf{l}'_2)$). For a given pairs of \mathbf{l}'_1 and \mathbf{l}'_2 , we choose the third vector \mathbf{l}'_3 from the condition $\mathbf{l}'_3 = -\mathbf{l}'_1 - \mathbf{l}'_2$ and then accept the triplet of $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$ (otherwise discard it) if the length l'_3 is in the range satisfies the condition $l_{\min} \leq l'_3 < l_{\max}$. Then we assign each set of three vectors, $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$, to one of the triangle configuration bins labelled by (l_1, l_2, l_3) by sorting l'_1, l'_2 and l'_3 in the ascending order. For any of the sets of three vectors chosen in this way, two of the three vectors are in the upper half plane of Fourier space, while the remaining one is in the lower half plane. Hence, we miss triangles with configurations for which two vectors are in the lower plane and the other in the upper. However, we can recover these triangles by just flipping the signs of all three vectors and obtain the same value of bispectrum, due to the symmetry $\tilde{\kappa}_{-\mathbf{l}} = \tilde{\kappa}_{\mathbf{l}}^*$, which comes from the real condition of the lensing field. Thus we need not count the latter cases, but just twice the number of actually counted triangles to obtain N_{trip} , the number of independent triplets for a survey area of 25 sq. degrees. For some of the following results, we will use the measured N_{trip} when computing the Gaussian error contributions to the bispectrum covariance (the first term in Eq. 18). Although Eq. (16) gives a good approximation to N_{trip} for the limit of $l_i \gg 1$, the N_{trip} directly estimated from the simulation properly takes into account the effect of discrete grids in the lensing map.

Using the table of three vectors $(\mathbf{l}'_1, \mathbf{l}'_2, \mathbf{l}'_3)$ obtained in the process above, we estimate the bispectrum by averaging $\text{Re}[\tilde{\kappa}_{\mathbf{l}'_1} \tilde{\kappa}_{\mathbf{l}'_2} \tilde{\kappa}_{\mathbf{l}'_3}]$ from each realization. Note that the value of $\tilde{\kappa}_{\mathbf{l}'_i}$ is taken from the field at the grid that has two coordinate components (l'_{ix}, l'_{iy}) , not from the field at the grid denoted by the arrow of the vector \mathbf{l}'_i in Fig. 3. Although we take the real part of $\tilde{\kappa}_{\mathbf{l}'_1} \tilde{\kappa}_{\mathbf{l}'_2} \tilde{\kappa}_{\mathbf{l}'_3}$ for the average, this is not essential, because the estimated bispectrum satisfies the real condition to a very good approximation after the average over many triangles.

The dimensions of the resulting covariance matrices are: 16×16 , 204×204 and 16×204 (or 204×16) for the power spectrum

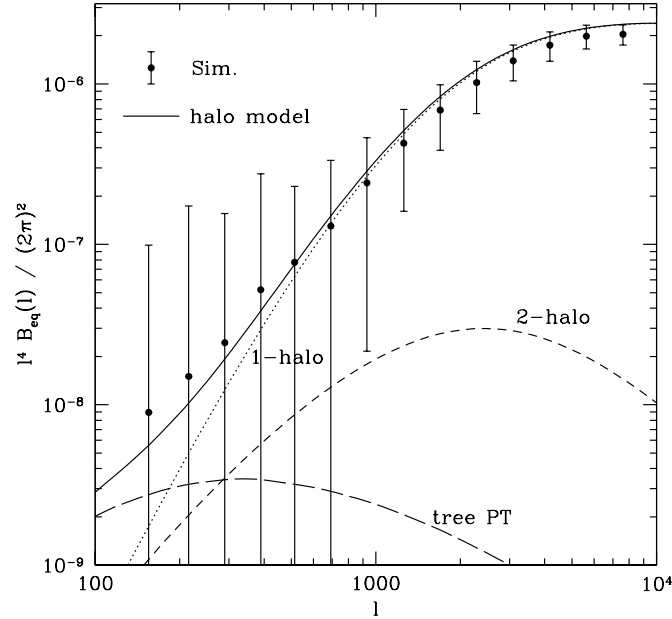


Figure 4. The lensing bispectrum for equilateral triangle configuration as a function of the side length l , where the multipole bins are logarithmically spaced by $\Delta \ln l = 0.3$. The data points with error bars are the bispectrum measured from the 1000 ray-tracing simulations with the source redshift $z_s = 1$. Each simulation covers an area of 25 deg^2 . The error bars show the scatters of the 1000 realizations, which therefore correspond to the measurement errors expected for the area of 25 deg^2 . The solid curve show the halo model prediction, while the dotted, short-dashed and long-dashed curves show the 1-, 2- and 3-halo term contributions to the bispectrum. For the 3-halo term, we used the tree-level perturbation theory prediction. Note that the bispectra for $l \gtrsim 4000$ may be affected by the resolution limit of the ray-tracing simulations.

covariance, the bispectrum covariance and the cross-covariance, respectively. The 1000 realizations are okay to compute the 204×204 elements of the covariance matrix (see Appendix of Takahashi et al. 2011), although a larger number of the realizations are ideally needed for a more accurate estimate of the covariance matrix. This is definitely true in a case that more triangle configurations are considered, e.g. as in the case of lensing tomography where different redshift slices are further needed to include (Takada & Jain 2004). Hence an analytical approach to compute the covariance matrices is useful.

4.3 Comparison of the simulation results and the halo model predictions

Fig. 4 plots an example of the measured bispectrum (*points* with error bars) and the halo model prediction (*lines*) for equilateral triangle configurations. The halo model fairly well agrees with the simulation results, although it under-estimates the bispectrum amplitude around $l \sim$ a few 100 and over-estimates at $l \gtrsim 1000$. An improvement of the halo model accuracy may be available by refining the halo model calculation, as done in Valageas et al. (2012). However, in this paper, we do not pursue this possibility, because the primary purpose of this paper is to study the bispectrum covariance and the information content of the lensing bispectrum, where the inaccuracy of the halo model predictions does not have a significant impact.

In Figs. 5 – 7, we study the covariance matrices of the lensing bispectra for some representative triangle configurations of the 204 triangles. Note that, in these results, the two triangle configurations of the covariance have the same length(s) for at least one side length (e.g., $l_1 = l'_1$). Hence the covariance terms $O(BB)$ and $O(PT)$ in Fig. 1 are not vanishing for the covariances shown in these figures. First, in Fig. 5, we study the diagonal terms of the bispectrum covariance matrix for equilateral triangles, $\text{Cov}[B_{eq}(l), B_{eq}(l)]$, as a function of the side length l . The points show the simulation results. The jagged feature at small l bins is due to the effect of discrete pixels in the lensing maps which causes the number of triangles to be non-smooth in the small l range. The different curves are the halo model predictions for the covariance matrix, which are computed based on Eqs. (18) and (19). The dotted curve shows the Gaussian error contribution that scales with $P(l)^3$, where we used the number of triangles directly computed from the simulated lensing map in computing the Gaussian term in Eq. (18). The long dot-dashed, short dot-dashed, long dashed and short dashed curves are the non-Gaussian terms in Eq. (18). For these calculations, we used the $P(l)$ and $B_{eq}(l)$ directly estimated from the 1000 ray-tracing simulations, while we used the halo model in § 3.5 to compute the higher-order functions, which are needed to compute the non-Gaussian terms. The solid curve is the total power of the covariance matrix, which is computed by summing all the different curves of the model predictions. The figure clearly shows that the non-Gaussian errors become significant at $l \gtrsim$ a few 100, and the model predictions are in good agreement with the simulation results, if including the HSV term. The HSV term is dominated over the other terms at $l \gtrsim 1000$. These findings are similar to the results in the power spectrum covariance (see Figs. 5, 6 and 7 in Sato et al. 2009).

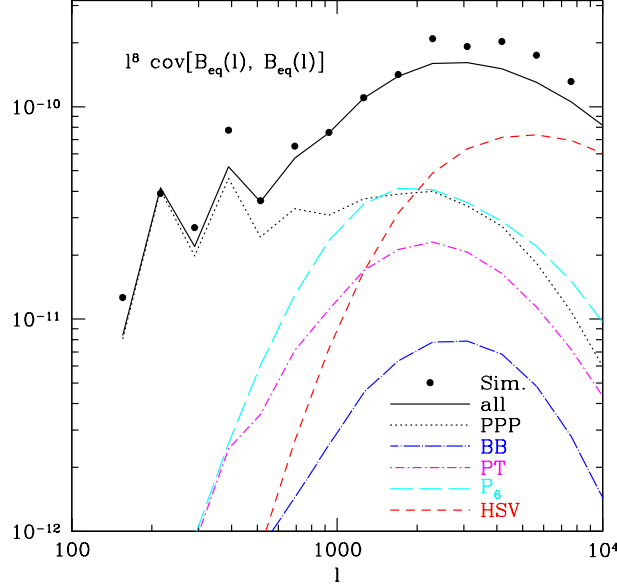


Figure 5. Diagonal terms of the bispectrum covariance matrix for equilateral triangle configurations as a function of the side length l , as in the previous figure. The covariance amplitude is shown in the unit of $l^6 \text{Cov}[B_{\text{eq}}(l), B_{\text{eq}}(l)]$, since $l^3 B_{\text{eq}}(l)$ gives the contribution to the skewness $\langle \kappa^3 \rangle$. The points are the simulation results estimated from the scatters of 1000 simulations. The other curves are the halo model predictions, which are computed based on the method described in § 3.2. The dotted, long dot-dashed, short dot-dashed, and long dashed curves are the contributions that are proportional to P^3 , B^2 , PT and P_6 , respectively (also see Fig. 1). The contributions of $O(P^3)$ are the Gaussian error contributions, and the other terms are the non-Gaussian pieces. For these calculations, we used the power spectra $P(l)$ and the bispectra $B(l)$ directly estimated from the simulations. The short dashed curve is the HSV contribution, which dominates over other terms at multipole bins, $l \gtrsim 2000$. The solid curve is the total contribution, the sum of all the terms.

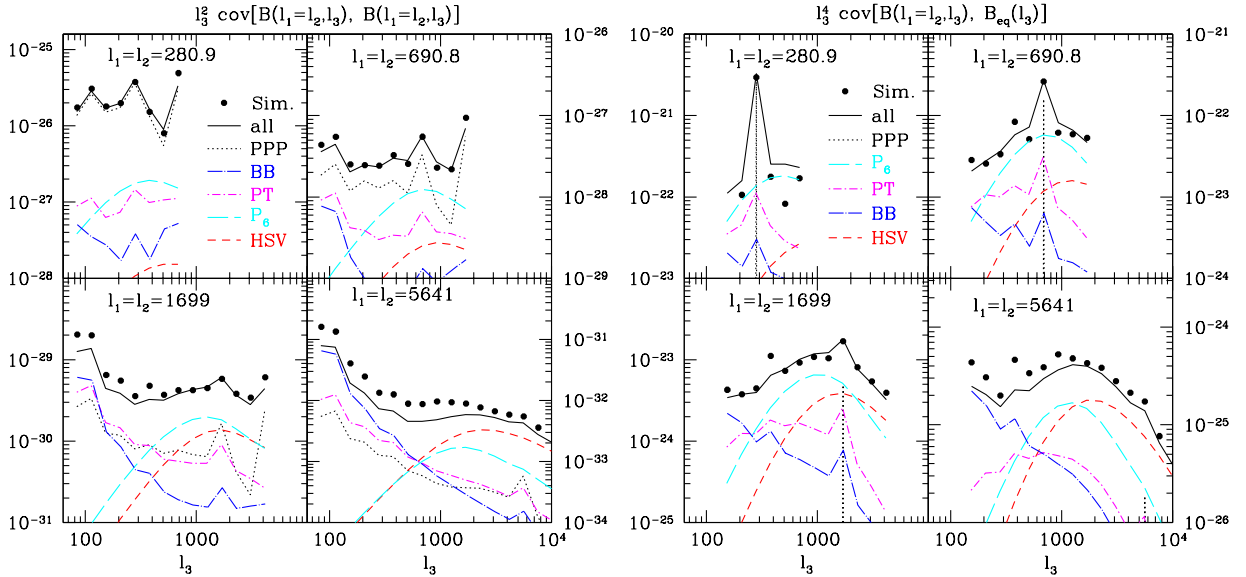


Figure 6. Similar to the previous plot, but for different triangle configurations. *Left panels:* The covariance matrix for isosceles triangle configurations, $B(l_1, l_2, l_3)$, with $l_1 = l_2$. The different panels are for different side lengths l_1 and l_2 ; within each panel, the covariance matrix is shown as a function of l_3 . Note that the covariance matrix is shown in the multipole range where the triangle conditions $|l_j - l_k| \leq l_i \leq l_j + l_k$ are satisfied, but the condition $l_1 \leq l_2 \leq l_3$ is not imposed. *Right panels:* The covariance matrix elements between the bispectra of isosceles and equilateral triangles. The Gaussian terms of P^3 denoted by the vertical dotted lines appear at a particular value of l_3 , where the two triangles become the same, i.e. equilateral triangles with $l_1 = l_2 = l_3$.

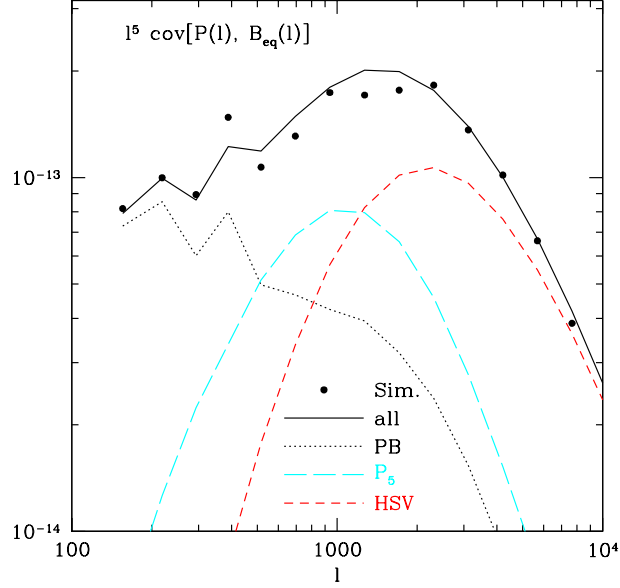


Figure 7. Cross-covariance between the power spectrum and the bispectrum of equilateral triangles, $\text{Cov}[P(l), B_{\text{eq}}(l)]$, as a function of l . The covariance amplitude is plotted in the units of $l^5 \text{Cov}[P, B_{\text{eq}}]$, because $l^2 P$ and $l^3 B$ contribute to $\langle \kappa^2 \rangle$ and $\langle \kappa^3 \rangle$, respectively. Note that there is no Gaussian error contribution to the cross-covariance, because it arises from the 5-point correlation functions. The different curves are the model predictions that are computed based on the method in § 3.3.

Fig. 6 shows the covariance matrices for isosceles triangle configurations (left panel) and the off-diagonal components between different triangle configurations, equilateral and isosceles triangles here. For this plot, we did not use the condition $l_1 \leq l_2 \leq l_3$ for presentation purpose, but all the triangles shown here are indeed included in the covariance matrix elements we will use in the following analysis. Even for these more general triangle configurations, the model predictions including the HSV effect are in good agreement with the simulations. One may find that the covariance amplitudes or some non-Gaussian terms peak at some particular value of l_3 in the x -axis. This happens when the isosceles triangles become to have higher symmetry, equilateral triangles; the number of independent triangles become smaller for higher-symmetry triangles, leading to the greater covariance amplitudes.

In Fig. 7, we study the cross-covariance between the power spectrum and the bispectrum of equilateral triangles, $\text{Cov}[P(l), B_{\text{eq}}(l)]$, as a function of the multipole bin l . There is no Gaussian error contribution to the cross-covariance, as it arises from the 5-point functions of the lensing field. The figure again shows that the model predictions including the HSV term can well reproduce the simulation results.

In Fig. 8, we compare the halo model predictions with the simulation results for all the matrix elements of the power spectrum covariance, the bispectrum covariance and the cross-covariance, in one figure. To do this, we use the correlation coefficients of the covariance matrices defined as

$$r_{ij}^{XY} \equiv \frac{\text{Cov}[X_i, Y_j]}{\sqrt{\text{Cov}[X_i, X_i] \text{Cov}[Y_j, Y_j]}}, \quad (25)$$

where X and Y are the power spectrum or the bispectrum, and the subscript i or j in X or Y denote the i -th multipole bin or the i -th triangle configuration; $X_i = P(l_i)$, $X_i = B(l_i)$, and so on. The diagonal components $r_{ii} = 1$ by definition. If $r_{ij} = 0$, it means no correlation between the spectra X_i and Y_j , while the higher values of r_{ij} mean stronger correlations. For illustration purpose, we use the following indices of the 204 triangles so that the different triangles are indexed in increasing order of l_3 :

$$\begin{aligned} \Delta(i_{l_1}, i_{l_2}, i_{l_3}) = & (1, 1, 1), \\ & (1, 1, 2), (1, 2, 2), (2, 2, 2), \\ & (1, 1, 3), (1, 2, 3), (1, 3, 3), (2, 2, 3), (2, 3, 3), (3, 3, 3), \\ & (1, 3, 4), \dots, \\ & \vdots \\ & (1, 16, 16), \dots, (14, 14, 16), (14, 15, 16), (14, 16, 16), (15, 15, 16), (15, 16, 16), (16, 16, 16), \end{aligned} \quad (26)$$

where we have used the 16 logarithmically-spaced multipole bins of l . Note that, for a given l_3 -bin, the other multipole bins (l_1, l_2) are listed in increasing order of l_1 ($l_1 \leq l_2$ for each triangle index). With this triangle index, the higher-index triangle configuration corresponds

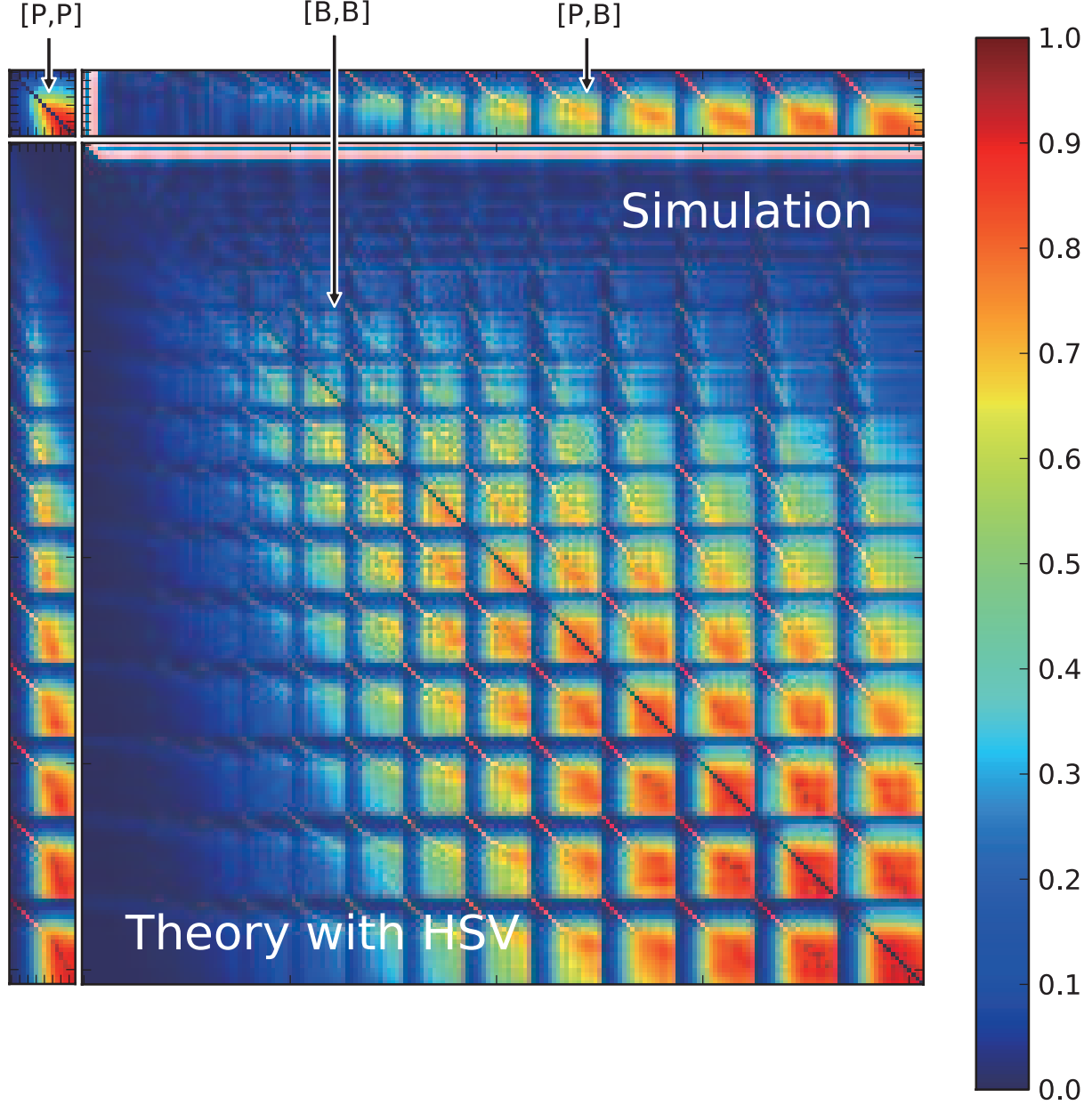


Figure 8. Correlation coefficients r_{ij}^{XY} (defined by Eq. 25) for the power spectrum and bispectrum covariance and the cross-covariance between the power spectrum and the bispectrum, where $X, Y = P$ or B and the indices i or j denote the multipole bin or the triangle configuration, e.g. $X_i = B(l_i)$. Note that the diagonal elements $r_{ii}^{XX} = 1$ by definition. The upper-right matrix elements show the simulation results, while the lower-left elements show the halo model predictions with the HSV effect. The upper-left square-shaped panel (16×16 elements) shows $r_i^{PP} j$ for the power spectrum covariance. The lower-right panel (204×204) shows the bispectrum covariance matrix r_{ij}^{BB} . The upper-right or lower-left rectangular-shape panels (16×204 or 204×16) show the cross-covariance matrix r_{ij}^{PB} or r_{ij}^{BP} . With increasing the multipole bins, the correlation coefficients have larger values and approach $r_{ij}^{XY} \simeq 1$.

to the triangles having higher multipoles or larger side lengths. The figure shows that the halo model well reproduces the two-dimensional features of the covariance matrices seen from the simulations¹⁰. Note that, for the bispectrum covariance, most of the off-diagonal terms have non-Gaussian contributions from the terms $O(P_6)$ (see Fig. 1) and the HSV effect, while the other terms all vanish, in contrary to Figs. 1 and 6, because the two triangles are in different shapes. The correlation coefficients become greater at higher multipoles, almost $r_{ij} \simeq 1$, meaning significant correlations between the different spectra (power spectra, bispectra and the cross-covariance).

For comparison, Fig. 9 shows the results if ignoring the HSV effect in the halo model predictions (the simulation results are the same to the previous figure). It is clear that the model predictions do not reproduce the simulation results.

¹⁰ Note that, in Fig. 8, we fully used the halo model to compute the covariance matrix elements including the power spectrum and bispectrum, unlike in Figs. 5, 6 and 7.

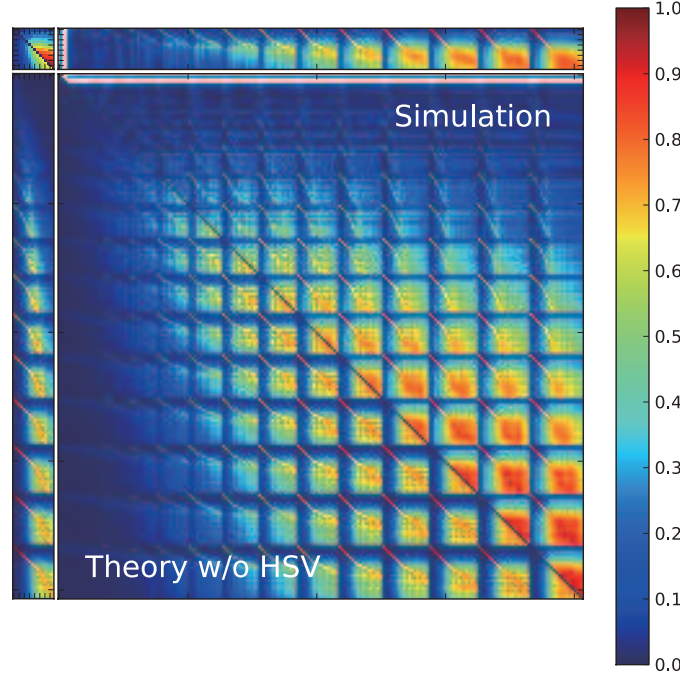


Figure 9. Similar plot to the previous figure, but the halo model predictions without the HSV effect are shown (the lower-left matrix elements). Note that the simulation results (the upper-right elements) are the same as in the previous figure. It is evident that the halo model predictions are well below the simulation measurements for large multipoles owing to the missing HSV terms.

4.4 Information content of the lensing bispectrum

As we have studied, the non-linear structure formation induces non-Gaussian errors in the weak lensing field, provoking significant correlations between the power spectra of different multipoles and the bispectra of different triangle configurations as well as significant cross-correlations between the power spectra and the bispectra. Then a more fundamental, important question arises: how much do the lensing bispectra carry additional information to the lensing power spectrum? Can joint measurements of the power spectra and bispectra recover the information content of the Gaussian field, which the primordial density field of structure formation should have had as in the CMB case? In this section, we address these questions.

A useful quantity to quantify the impact of the non-Gaussian errors is the expected signal-to-noise ratio (S/N) for a measurement of the lensing power spectra and bispectra for a given survey that is characterized by its area and shot noise parameters. The S/N is sometimes called the *information content* (Tegmark et al. 1997) (also see Takada & Jain 2009, and references therein). For the power spectrum measurement, the S/N is defined as

$$\left(\frac{S}{N}\right)_P^2 \equiv \sum_{l_i, l_j < l_{\max}} P(l_i) [C^P]_{ij}^{-1} P(l_j), \quad (27)$$

where the summation indices i, j run over multipole bin indices up to a given maximum multipole l_{\max} , and $[C^P]^{-1}$ is the inverse of the power spectrum covariance matrix (Eq. 11 for the expression). The S/N is equivalent to a precision of determining the log of the power spectrum amplitude parameter, $\ln A$, from the power spectrum measurement up to a given maximum multipole l_{\max} , assuming that the shape of the lensing power spectrum is perfectly known. The S/N is independent of the multipole bin width as long as the bin width is fine enough to capture the shape of the lensing power spectrum (on the other hand, recall that the relative strength of the non-Gaussian errors to the Gaussian errors in each covariance matrix element depends on the multipole bin width assumed).

Similarly, the S/N for the bispectrum measurement or the information content about the bispectrum amplitude parameter is defined as

$$\left(\frac{S}{N}\right)_B^2 = \sum_{\{l_i\}, \{l_j\} \leq l_{\max}} B_i [C^B]_{ij}^{-1} B_j, \quad (28)$$

where the summation indices i, j run over triangle configurations, and we include all the triangle configurations whose side lengths are up to a given maximum multipole l_{\max} . The $(S/N)_B$ gives an estimate on the precision of constraining the amplitude parameter of lensing bispectra, assuming that the dependence on triangle configurations is perfectly known. Again, as long as the multipole bin width or the steps between different triangle configurations are fine enough to capture the dependence of the lensing bispectra on triangle configurations, the $(S/N)_B$ is independent of the multipole bin width.

We also consider the S/N for a combined measurement of the lensing power spectra and bispectra up to a given l_{\max} . In the presence of

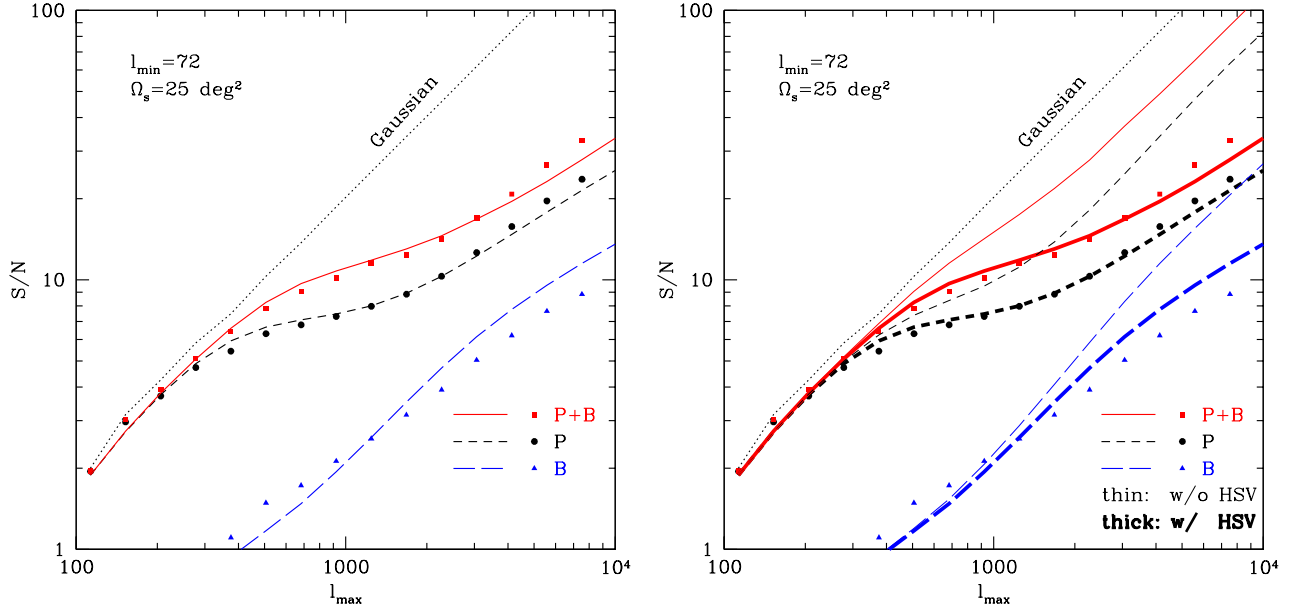


Figure 10. *Left panel:* Cumulative signal-to-noise ratios (S/N) for the power spectrum (P), the bispectrum (B) and the joint measurement ($P + B$) for a survey area of 25 deg^2 and source redshift $z_s = 1$. It is shown as a function of the maximum multipole l_{max} , where the power spectrum and/or bispectrum information are included over $l_{\text{min}} \leq l \leq l_{\text{max}}$ (see Eqs. 27, 28 and 31). Note that we set $l_{\text{min}} = 72$ and did not include the shape noise contamination here – it is shown in the next figure. The circle, triangle and square symbols are the simulation results, computed from the 1000 realizations, for P , B and $P + B$ measurements, respectively. The short-dashed, long-dashed and solid curves are the halo model predictions. Adding the bispectrum to the power spectrum does increase the S/N amplitude, e.g. by about 50% at $l_{\text{max}} \simeq 10^3$. For comparison, the thin dotted curve shows the S/N for the power spectrum for the Gaussian field, which the primordial density field should have contained. *Right panel:* The thinner curves are added to the left panel to show the model predictions without the HSV contribution. For $l > 1000$, the HSV contribution lowers the S/N significantly.

the non-Gaussian errors, the total S/N is *not* simply a sum of the S/N 's of the power spectra and the bispectra due to the cross-covariance, because the two spectra are not independent. To study this, we first define the data vector for the joint measurement as

$$\mathbf{D} = \{P_1, P_2, \dots, P_{n_b}, B_1, B_2, \dots, B_{i_{\text{triang}, \text{max}}}\}. \quad (29)$$

The covariance matrix for the data vector \mathbf{D} is given as

$$\mathbf{C}^{P+B} = \begin{pmatrix} \mathbf{C}^P & \mathbf{C}^{PB} \\ \mathbf{C}^{PB} & \mathbf{C}^B \end{pmatrix}, \quad (30)$$

where the \mathbf{C}^{PB} is the cross-covariance between the power spectrum and the bispectrum. Then, the total S/N for the combined measurement is similarly defined as

$$\left(\frac{S}{N}\right)_{P+B}^2 = \sum_{i,j \leq l_{\text{max}}} D_i [\mathbf{C}^{P+B}]_{ij}^{-1} D_j. \quad (31)$$

Fig. 10 shows the expected S/N 's for measurements of the power spectra and the bispectra for a survey area of 25 square degrees (i.e. the area of the ray-tracing simulation), as a function of the maximum multipole l_{max} up to which the power spectrum and/or bispectrum information are included. Note that the minimum multipole is fixed to $l_{\text{min}} = 72$, and we did not include the shot noise contamination to the error covariance matrices, so the results solely correspond to the cosmological information contents. The circle, triangle and square symbols are the simulation results for the S/N 's of the power spectra, the bispectra and the joint measurements, respectively, which are computed using the 1000 realizations. The thick short-dashed, long-dashed and solid curves are the halo model predictions. First of all, the lensing bispectra add a new information content to the power spectrum measurement. To be more quantitative, adding the bispectrum measurement increases the S/N values by about 50% for $l_{\text{max}} \simeq 10^3$ compared to the power spectrum measurement alone, where $l_{\text{max}} \simeq 10^3$ or a few 10^3 are the target maximum multipoles for the upcoming weak lens surveys. This improvement is equivalent to about 2.3 larger survey area for the power spectrum measurement alone; that is, the same data sets can be used to obtain the additional information, if the bispectrum measurement is combined with the power spectrum measurement. Secondly, the halo model predictions are in nice agreement with the simulation results. Note that the total S/N for the joint measurement ($P + B$) is close to the linear sum of the S/N values $((S/N)_P$ and $(S/N)_B$), not the sum of their squared values $(S/N)^2$, due to the significant cross-covariance between P and B (see Appendix C in Takada & Bridle 2007, for the similar discussion). If ignoring the cross-covariance, adding the bispectrum measurement does not much

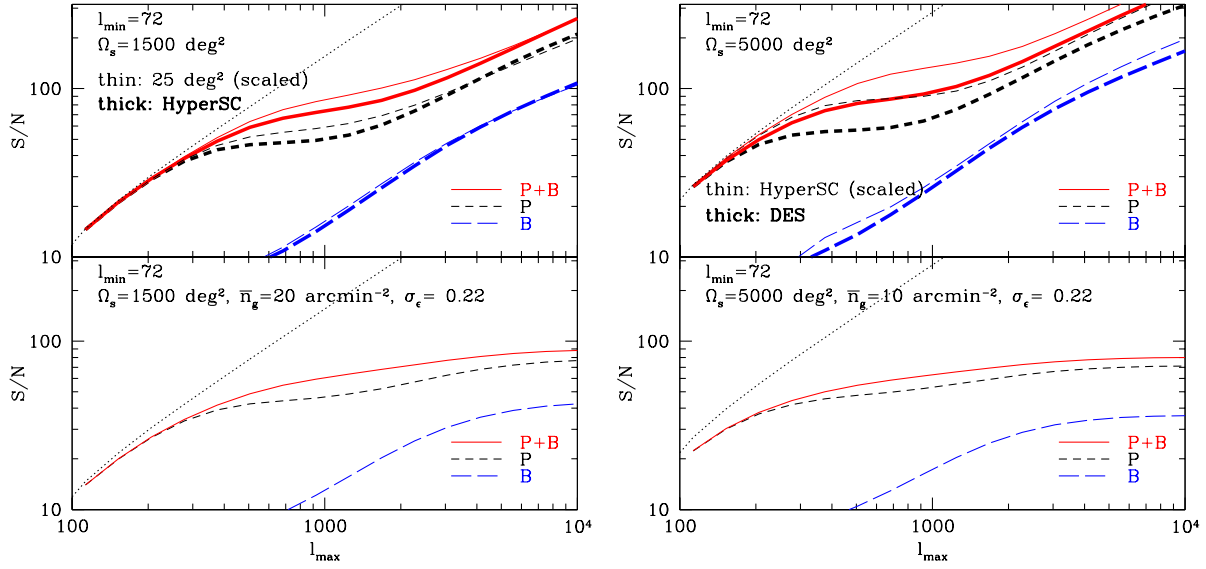


Figure 11. Cumulative S/N with the HSV effect, as in the previous figure, but here showing the forecasts for upcoming weak lensing surveys, the Subaru Hyper Suprime-Cam (HyperSC) survey and the Dark Energy Survey (DES) in the left and right panels, respectively. These surveys are characterized by the survey parameters: survey area (Ω_s), mean source redshift (\bar{z}_s), and mean (effective) number density of source galaxies (\bar{n}_g). We assumed $\Omega_s = 1500$ sq. degrees, $\bar{z}_s = 1$, and $\bar{n}_g = 20 \text{ arcmin}^{-2}$ for the HyperSC survey and $\Omega_s = 5000$ sq. degrees, $\bar{z}_s = 0.7$, and $\bar{n}_g = 10 \text{ arcmin}^{-2}$ for the DES survey. We set $\sigma_\epsilon = 0.22$ for the rms intrinsic ellipticity per component for both the surveys. The upper and lower plots in each panel show the results without and with the shot noise contamination. The thin curves in the upper plot of the left panel shows the S/N values obtained by scaling the results for 25 sq. degrees in Fig. 10 assuming $S/N \propto \Omega_s^{1/2}$. The thin curves in the upper plot of the right panel is similar, but obtained by scaling the HyperSC results in the left panel to the results for 5000 sq. degrees assuming $S/N \propto \Omega_s^{1/2}$ and the source redshift $\bar{z}_s = 1$. The lower plots in each panel show the results including the shot noise contribution to the covariance. Clearly shot noise removes information at high l and flattens the S/N curves above $l > 1000$.

improve the S/N value (only by 5% or so). Hence it is important to take into account how the power spectrum and the bispectrum are correlated for the underlying true cosmology.

For comparison, the dashed curve shows the S/N values for the Gaussian case, which is the full information content of the lensing field if the lensing field is a random Gaussian field as in the CMB case or the primordial fluctuation fields. The figure clearly shows that, even if the lensing field originates from the primordial Gaussian field, the joint power spectrum and bispectrum measurement does not recover the full information content of the Gaussian field after the nonlinear structure formation processes. This implies that the higher-order functions beyond the bispectra are also important to recover the information content. Or the Gaussian information content or the initial memory can not be fully recovered; some of the information content may be lost after the nonlinear structure formation. However, one caution is needed to interpret the results in Fig. 10. The thin curves in the figure show the halo model predictions if ignoring the HSV effect in the covariance matrices. This is the case that there is no fluctuations of scales outside the lensing field of 25 deg^2 ; or the Fourier modes in the region obey periodic boundary conditions even if the field is highly non-Gaussian, although this does not hold for our Universe. In this case, the S/N values are about 25, 11, 18 for the Gaussian case, the power spectrum (P) and the joint measurement ($P + B$), respectively, for $l_{\text{max}} \simeq 10^3$ (more exactly the bin of $l_{\text{max}} = 1245$). Thus, the power spectrum contains about 44% of the Gaussian information, and the joint measurement can recover about 72% of the information, almost factor of 2. Thus the small S/N value for $P + B$ in Fig. 10, compared to the Gaussian case, is mostly due to the HSV effect. As discussed in § 3.4, the HSV effect alters the overall amplitude of the power spectrum or bispectrum, and preserve their shapes. Hence the HSV effect may give the worst case degradation for the amplitude parameter, but may not cause any serious degradation in parameters that are sensitive to the shapes of the lensing spectra than naively thought. Thus a genuine impact of the HSV effect on the parameter estimation needs to be further studied and this is our future work.

In Fig. 11 and Table 1, we show the S/N 's expected for the upcoming wide-field weak lensing surveys, the Subaru Hyper Suprime-Cam (HyperSC) survey and the Dark Energy Survey (DES), which are characterized by survey area, the mean source redshift and the mean number density of source galaxies: $\Omega_s = 1500$ sq. degrees, $\bar{z}_s = 1$ and $\bar{n}_g = 20 \text{ arcmin}^{-2}$ for the HyperSC survey, while $\Omega_s = 5000$ sq. degrees, $\bar{z}_s = 0.7$ and $\bar{n}_g = 10 \text{ arcmin}^{-2}$ for the DES, respectively. Here we employed the halo model to compute the S/N 's, and assume a circular survey geometry for simplicity; the survey area is given as $\Omega_s = \pi\Theta_s^2$. The figure and table show that the upcoming surveys promise to allow a significant detection of the lensing bispectrum; $(S/N) \simeq 26$ or 29 for the HyperSC or DES surveys, respectively, when assuming $l_{\text{max}} \simeq 2000$ and including the shot noise effect. This means that the theoretical predictions for the lensing bispectra need to be accurate to a prediction of a few % in the fractional errors for the upcoming surveys. It can be also found that the bispectra add the new information, increasing the total S/N by about 20 – 30% compared to the power spectrum alone. This is equivalent to a factor of 1.4 – 1.7 larger survey area.

Expected cumulative S/N 's for the upcoming weak lensing surveys

Survey $l_{\max} (\simeq)$ shot noise (σ_ϵ)	Subaru HyperSC Survey				DES			
	1000 w/o	1000 with	2000 w/o	2000 with	1000 w/o	1000 with	2000 w/o	2000 with
$(S/N)_P$	53	48	74	57	75	56	116	63
$(S/N)_B$	19	16	35	26	33	20	59	29
$(S/N)_{P+B}$	78 (48%)	64 (33%)	98 (32%)	72 (26%)	103 (37%)	66 (18%)	145 (25%)	73 (16%)

Table 1. Cumulative S/N 's of the power spectrum (P), the bispectrum (B) and the joint measurement ($P + B$), expected for the upcoming weak lensing surveys, the Subaru Hyper Suprime-Cam survey (HyperSC) and the Dark Energy Survey (DES), as in Fig. 11. Here we consider $l_{\max} \simeq 1000$ and 2000 for the maximum multipole (more exactly, the bins of $l_{\max} = 1245$ and 2268). The column denoted by “w/o” or “with” in the row “shot noise” gives the S/N values with and without the intrinsic shape noise contribution to the covariances. The percentage in the parenthesis shows the improvement of S/N value for the joint measurement ($P + B$) compared to the power spectrum alone (P).

Fig. 11 also shows that the HSV effect is significant for the upcoming surveys. The thin curves in the upper plots of the left-panel are the S/N 's values computed by scaling the S/N 's values for 25 deg^2 in Fig. 10 assuming $S/N \propto \Omega_s^{1/2}$. Since the covariance terms besides the HSV term scale as $1/\Omega_s$, the differences between the thick and thin curves are due to the HSV term which depends on a survey area via the shape of the matter power spectrum convolved with the survey window function (see discussion below Eq. 14). It can be found that the S/N values are smaller than the naive scaling results, meaning that the HSV effect has a more slowly-decreasing function than $1/\Omega_s$. The upper plot of the right panel shows the similar plot, but for the DES results obtained by scaling the S/N values of the HyperSC results in the left panel. The S/N values are smaller than the naive expectation due to the following reasons. The DES has a typical source redshift of $\bar{z}_s = 0.7$ compared to $\bar{z}_s = 1$ for the HyperSC survey, and therefore is more sensitive to large-scale structures at lower redshifts, which are more evolving and are in more nonlinear regime. Thus the S/N values for the DES are more degraded by the non-Gaussian errors.

The lower panels in Fig. 11 show the results when further including the contribution of shot noise (intrinsic shape noise) to the covariance matrices. The shot noise leads to a saturation of the S/N values of the *cosmological* signals. However, the effects of the non-Gaussian errors and the HSV contributions are important around $l_{\max} \sim 10^3$, the target maximum multipole for the planned surveys, because the systematic effects such as the highly nonlinear clustering effect and/or the baryonic effect become more significant at the higher multipoles (White 2004; Zhan & Knox 2004; Huterer & Takada 2005; Huterer et al. 2006).

4.5 Principal component analysis of the lensing covariance matrices

A principal component analysis (PCA) of the power spectrum and bispectrum covariance matrices is useful to quantify how the different power spectra and/or bispectra are correlated with each other and how many independent modes or triangles contribute to most of the information contents (Takada & Jain 2009) (also see Scoccimarro 2000, for the 3D bispectrum case). Since the covariance matrix is symmetric by definition, it can always be decomposed as

$$C_{ij}^X = \sum_a S_{ai}^X (\lambda_a^X)^2 S_{aj}^X, \quad (32)$$

where X is the lensing power spectrum or bispectrum, λ_a^X is the a -th eigenvalue or principal component, $(\mathbf{S}^X)^{-1} = (\mathbf{S}^X)^T$, $\sum_k S_{ik}^X S_{jk}^X = \delta_{ij}^K$ and \mathbf{S}^X is normalized so as to satisfy $\sum_j (S_{ij}^X)^2 = 1$. The matrix S_{ai}^X is considered as the projection matrix as it describes how the power in the i -th multipole bin or triangle configuration is projected onto the a -th eigenmode. Using this representation, the inverse of the covariance matrix is given by $[\mathbf{C}^X]_{ij}^{-1} = S_{ia}^X (1/\lambda_a^X)^2 S_{ja}^X$. Hence, the S/N values (Eqs. 27 or 28) can be rewritten as

$$\left(\frac{S}{N}\right)_X^2 = \sum_a \left\{ \frac{1}{\lambda_a^X} \sum_i S_{ai}^X X_i \right\}^2, \quad (33)$$

where X_i is either $P(l_i)$, $B(l_i)$ or the joint ($P + B$). Thus, since $(1/\lambda_a^X) \sum_i S_{ai}^X X_i$ can be considered as the S/N for the a -th eigenmode, the above equation expresses the total S/N (for a given l_{\max}) as a sum of contributions from the independent eigenmodes. We can then re-order the eigenmodes such that the lower eigenmode has the higher $(S/N)^2$; e.g., the first eigenmode has the highest $(S/N)^2$.

Fig. 12 shows the PCA results for the power spectrum measurement for the Subaru HyperSC-type survey including the shot noise effect as in Fig. 11. This figure can be compared with Fig. 5 in Takada & Jain (2009), where the HSV effect was not included. The short-dashed and dotted curves show the differential $(S/N)^2$ values of each i -th multipole bins with and without the HSV effect, showing how the total $(S/N)^2$ value increases with adding the i -th multipole bin from the $(S/N)^2$ up to the $(i - 1)$ -th bin. It should be stressed that the power spectrum of $l \sim 10^3$ has a local minimum, due to the significant HSV contribution.

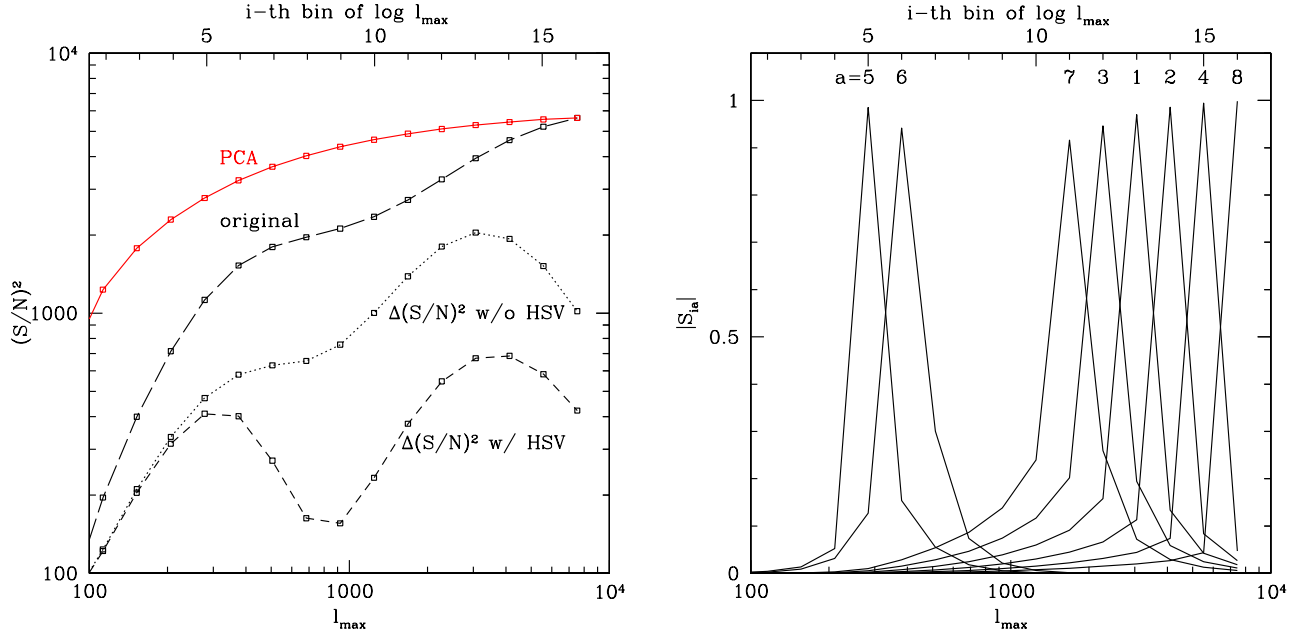


Figure 12. Principal component analysis (PCA) for the power spectrum covariance for the 16 logarithmically-spaced multipole bins over $72 \leq l \leq 8748.8$, for the Subaru HyperSC-type survey with 1500 deg^2 and with the shot noise. *Left panel:* The long-dashed curve shows the cumulative $(S/N)^2$ value for the power spectrum as in the left-lower panel of Fig. 11, but for $(S/N)^2$ instead of S/N (the points denote the central value of each multipole bin). The short-dashed and dotted curves show the differential contribution to the $(S/N)^2$ value at each multipole bin, with and without the HSV effect, respectively. The solid curve shows how the $(S/N)^2$ value is recovered by adding the PCA eigenmodes, where the PCA eigenmodes are ranked in decreasing order of the differential $(S/N)^2$ value (see Eq. 33), i.e. the higher $(S/N)^2$ for the lower eigenmode. Including up to about 7 eigenmodes, less than half of the 16 multipole bins, can recover about 90% of the total $(S/N)^2$ value. *Right panel:* The projection matrix $|S_{ia}|$ for the first 8 eigenmodes, where the index a denotes the a -th eigenmode. The projection matrix for each eigenmode peaks at some multipole bin, but has tails to different multipole bins. In particular, the eigenmodes around a few 10^3 have longer tails, reflecting significant correlations between different multipole bins due to the non-Gaussian errors.

The top dashed curve shows the total $(S/N)^2$ values as a function of the maximum multipole l_{max} , up to $l_{\text{max}} = 8748.8$ (the 16-th multipole bin). On the other hand, the solid curve shows how the cumulative $(S/N)^2$ value increases by adding the a -th eigenmode (Eq. 33)¹¹. The figure shows that, among the 16 multipole bins, including about 7 eigenmodes (about half of the multipole bins) can recover about 90% of the total $(S/N)^2$. The other eigenmodes are relatively less important due to strong correlations between the different multipole bins. The right panel shows the projection matrix $|S_{ia}|$ for each multipole bin, showing how the neighboring multipole bins are correlated with each other. In particular, the power spectra around $l \simeq 10^3$ show significant correlations.

Fig. 13 shows the PCA analysis for the bispectrum covariance matrix, where we included the 204 triangles over $72 \leq l \leq 8748.8$. First, it can be found that the different triangle configurations contribute to the total $(S/N)^2$ in a complex way. We use the triangle index, given by Eq. (26), and the higher-index triangle corresponds to triangle with larger side length of “ l_3 ” value, under the condition $l_1 \leq l_2 \leq l_3$. However, for a given l_3 bin, either of the two side lengths l_1 or l_2 can be much smaller than l_3 , which generally has the smaller differential $(S/N)^2$. As a result, the differential $(S/N)^2$ curve has jagged features. The solid curve shows the cumulative $(S/N)^2$ by adding the new PCA eigenmodes, and manifests that about 70 eigenmodes, about one third of the 204 triangles, carry about 90% of the total $(S/N)^2$.

Fig. 14 shows which triangle configurations contribute to the first 6 PCA eigenfunctions, which therefore have the 6 highest differential $(S/N)^2$ values. More exactly, each panel shows the projection matrix $|S_{ia}|$ for the a -th eigenmode ($a = 1, 2, \dots, 6$). In most panels, the projection matrix $|S_{ia}|$ has a single or a few peaks, reflecting that the higher- S/N eigenmode arises mostly from a single or a few different triangles. For example, the highest $(S/N)^2$ eigenfunction arises from isosceles triangles, which have side lengths of a few 10^3 (11- and 14-th multipole bins). However, note that we have used rather wide, logarithmically-spaced multipole bins, and most of triangle configurations available from Fourier space are close to isosceles triangle configurations. Hence, an exact shape of triangles giving most contributions to the total $(S/N)^2$ value slightly changes with the bin width. If we take a finer multipole bin, the triangles with the highest $(S/N)^2$ would differ from isosceles triangles.

¹¹ First, we made the principal component analysis of the power spectrum covariance matrix including up to the multipole bin $l = 8748.8$ (i.e. 16×16 matrix). Then, we re-ordered the eigenmodes in increasing order of the differential $(S/N)^2$ value. The solid curve in Fig. 12 shows how the cumulative S/N value increases by adding a new a -th eigenmode.

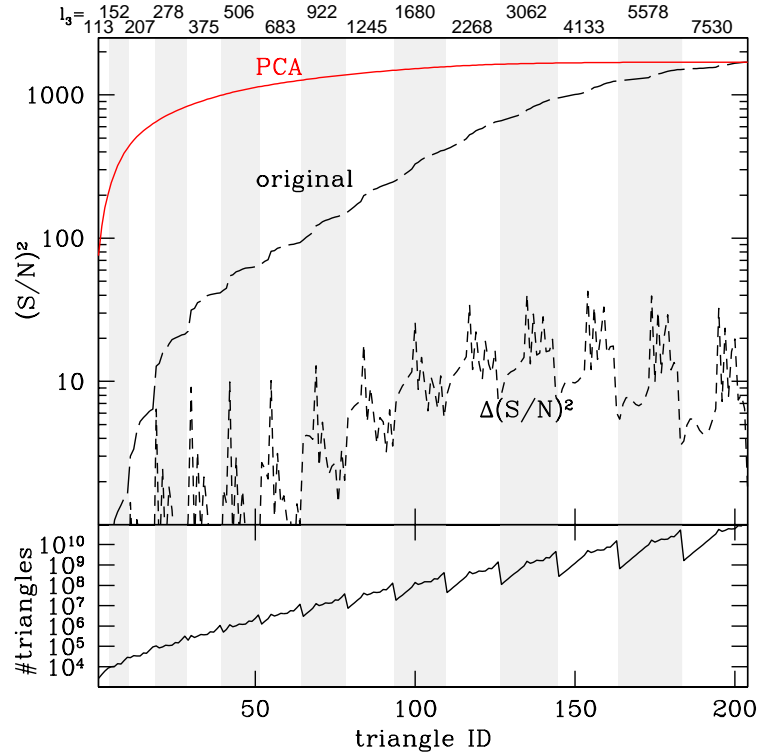


Figure 13. PCA analysis for the bispectrum covariance, for the 204 triangle configurations over $72 \leq l \leq 8748.8$, for the HyperSC-type survey as in the previous figure. The x -axis denotes either the triangle index or the PCA eigenmodes. The triangle indices are in the order given by Eq. (26); the larger indices correspond to triangles with the larger side length of l_3 . Note that the triangle parameters are (l_1, l_2, l_3) under the condition $l_1 \leq l_2 \leq l_3$. The vertical shaded regions denote the triangles with the same l_3 bin, as indicated by the top label of “ l_3 ”-bin values. *Upper panel:* The short-dashed curve shows the differential $(S/N)^2$ of each triangle configuration, while the long dashed curve shows the cumulative $(S/N)^2$ as in the left-lower panel of Fig. 11, but for the finer multipole bins (stepped by each triangle configuration). The solid curve shows the cumulative $(S/N)^2$ as one adds the different PCA eigenmodes. About 70 eigenmodes out of 204 triangles can recover 90% of the total $(S/N)^2$ of the bispectrum measurement over $72 \leq l \leq 8748.8$. *Lower panel:* The solid curve shows the number of independent triangles available from Fourier modes for the HyperSC-type survey (1500 square degrees), for each triangle configuration in the x -axis.

5 CONCLUSION AND DISCUSSION

We have studied the covariance matrices of the lensing power spectrum and the bispectrum, by using both 1000 ray-tracing simulations and the analytical halo model. We have found that there are significant non-Gaussian error contributions to the lensing covariance matrices; the power spectra or the bispectra at high multipoles of \gtrsim a few 10^2 are highly correlated with each other. In particular, we have shown that the mass density fluctuation modes of scales comparable with or outside a survey region cause significant non-Gaussian error contributions, which we call the halo sample variance (HSV) effect and have not been fully studied in the literature. With the HSV contributions, the halo model predictions accurately reproduce the covariance matrices measured from the 1000 simulations (Figs. 4–8).

Then we addressed how the lensing bispectrum carries additional information compared to the power spectrum, by including all the triangle configurations available up to the maximum multipole l_{\max} . Adding the bispectrum measurement improves the cumulative S/N values by about 20–50% compared to the power spectrum measurement alone, at a few 10^3 for l_{\max} . We also consider the cases with and without shape noise (Figs. 10–11 and Table 1). The improvement in S/N is equivalent to about 1.4 – 2.3 larger survey area for the power spectrum measurement alone. Hence, our results show that the same imaging data can be used to improve the constraining power of cosmological parameters, if we can combine the power spectrum and bispectrum measurements. We also find that the HSV effect is significant, leading to a large degradation in the cumulative S/N values. By using a principal component analysis of the covariance matrices, we find that about 1/3 eigenmodes of all the triangle configurations over the range of $10^2 \lesssim l \lesssim 10^4$ (70 eigenmodes compared to the 204 triangles for the multipole binning we assumed) carry most of the total information for the bispectrum measurement (Fig. 13). Thus, our results give a quantitative answer to long-standing question of how the bispectrum can be useful and complementary to the power spectrum while fully taking into account non-Gaussian errors and all triangle configurations.

Future surveys such as the Subaru HyperSC or DES surveys allow a significant detection of the lensing bispectra (Fig. 11 and Table 1). However, there are some practical issues for making the bispectrum measurement feasible for future surveys. First, we need to take into account effects of survey geometry and masking regions such as those due to saturated bright stars. This can be done by extending the

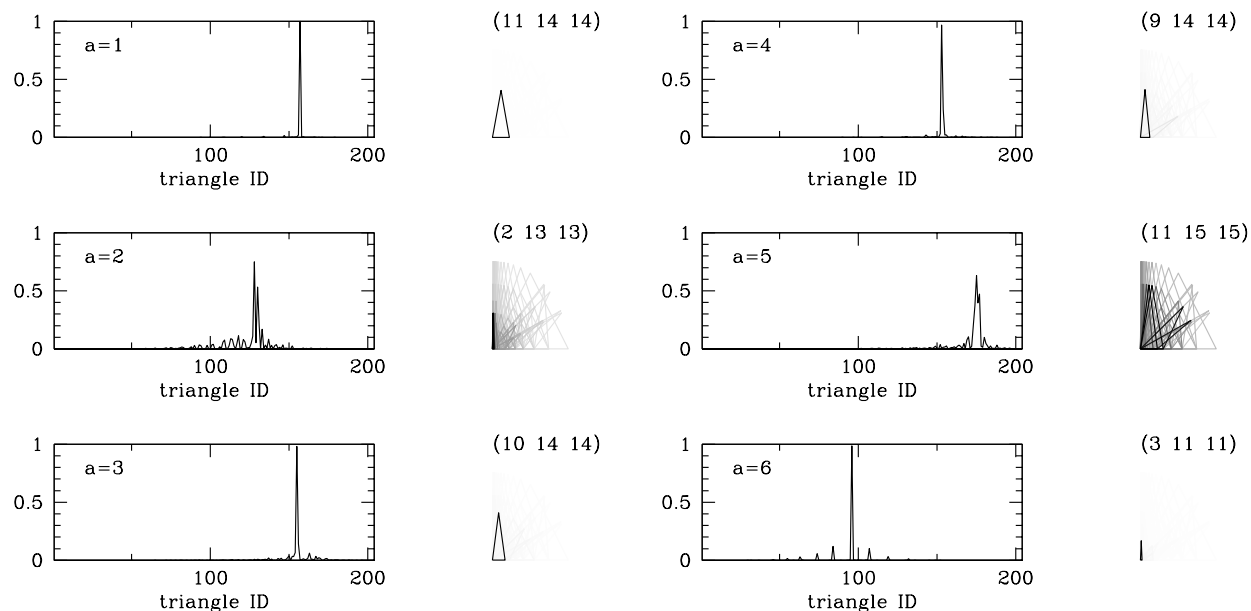


Figure 14. The curve in each panel shows which triangle configurations contribute to each PCA eigenmode for the bispectrum measurement in Fig. 13. We considered the first 6 PCA eigenmodes: the differential $(S/N)^2$ values are higher in order of the left-top, left-middle, ..., right-middle and right-bottom panels. More exactly, the absolute values of $|S_{ia}|$ are plotted, where S_{ia} is the projection matrix of the a -th eigenfunction onto the i -th triangle configuration. The triangle configurations shown on the right side of each panel are the triangle configurations with $S_{ia} \neq 0$, where the triangle sizes are plotted on the linear scale and the thickness of the lines is proportional to the amplitude of $|S_{ia}|$. The configurations with the largest contribution are labelled in the side length bins (l_1, l_2, l_3) at top. Note that we are using logarithmically-spacing multipole bins in this paper, so about 3/4 triangles available in the Fourier space are isosceles triangle configurations.

method in Hikage et al. (2011) for the power spectrum measurement to the bispectrum case, but has not yet been fully addressed in the literature. It is also important to explore a method of cleanly decomposing three-point correlations of E/B modes in the presence of the survey window function. An alternative approach is to measure the real-space three-point correlation functions, rather than the bispectrum, which does not suffer from the survey effects. However, in this case, there are even more significant correlations between the three-point correlations of different triangles, and an estimation of the covariance matrices involve multi-dimensional integrations of the bispectrum covariance, if one wants to use the halo model approach. Or a sufficiently large number of simulations may be used, but again it is not feasible to construct the covariance matrices for different cosmological models, which will be needed for parameter estimation. Hence, whether Fourier- or real-space is more useful/tractable for the three-point correlation measurements is still an open issue.

Although we focused on the cumulative S/N as a measure of the information content, S/N we defined is just one measure to quantify the complementarity of the lensing bispectrum. Specifically, the S/N quantifies the expected precision for determining the amplitude parameter, assuming that the shapes of the bispectrum or the power spectrum are perfectly known. Hence, the S/N is not necessarily a good measure to quantify the power of the bispectra for constraining cosmological parameters that are sensitive to the shape of the bispectrum (also see Takada & Jain 2009, for the similar discussion). Here we would like to stress a possible good news. Our results for the cumulative $(S/N)^2$ can be compared with Fig. 5 in Takada & Jain (2004), where only the Gaussian error contributions to the bispectrum covariance were included. The relative $(S/N)^2$ values between the power spectrum and bispectrum measurements are similar to the results shown in this paper, because the HSV effect degrades both the power spectra and the bispectra similarly. Even if the bispectrum S/N is smaller than that of the power spectrum by a factor of a few for $l_{\max} \simeq \text{a few } 10^3$, Takada & Jain (2004) showed that adding the bispectrum measurement can significantly improve accuracies of parameter estimation from the power spectrum measurement alone, by efficiently breaking parameter degeneracies. Thus we can expect similar improvement even with the inclusion of non-Gaussian errors. Exploring the genuine usefulness of the lensing bispectrum for cosmology is very interesting – the full forecasts for various cosmological parameters will be presented in future work.

In particular, to do such parameter forecasts, it would be much more interesting and useful if including tomographic information of the lensing field that is available from photometric redshift information of every source galaxies. It has been found that the lensing tomography can significantly improve the constraining power by recovering the radial-direction information of the lensing field (e.g. Takada & Jain 2004). However, for the lensing bispectrum tomography, we need to further include all the triangle configurations in multipole and redshift space in order to estimate the full potential. For example, if we consider three redshift bin tomography, the total number of triangles is $3^3 \times 204 = 5508$ for the same multipole binning as used in this paper. Thus, even 1000 realizations are not enough to reliably estimate the covariance matrices of the lensing bispectra. Hence we believe that the analytical formula we developed in this paper are essential to address these issues.

Perhaps as important as the improvement in statistical precision, the bispectrum enables self-calibration of systematic errors inherent

in the lensing tomography measurements such as photometric redshift errors and imperfect galaxy shape measurement. Lensing bispectra depend on systematic errors in a different way from the power spectrum, but the two spectra share the same large-scale structure, hence the same cosmological parameters that describe the underlying true cosmology. Thus combining the power spectrum and bispectrum measurements can be used to self-calibrate systematic errors and improve cosmological constraints (Huterer et al. 2006). Again, by fully taking into account the non-Gaussian errors, one needs to quantify how well self-calibration works for upcoming lensing surveys.

Planned multi-colour imaging surveys can also be used to measure other cosmological probes such as the abundance of galaxy clusters and the correlation functions of galaxies. Recently, Oguri & Takada (2011) proposed a new method of using the halo-shear correlations or the stacked lensing signals around massive halos as a function of the cluster redshifts, and showed that the stacked lensing tomography allows to constrain cosmological measurements to a high precision. In halo (cluster)-shear correlations, the signals arise from the large-scale structure at a particular redshift, the cluster redshift, which gives us a measure of the three-dimensional power spectrum, rather than the projected, angular power spectrum. As we have shown, massive clusters are key observables to understanding the HSV effect on the lensing power spectra, and it would be very interesting to pursue how we can combine the lensing power spectrum and bispectrum measurements with the cluster observables to improve cosmology parameter estimation by calibrating the HSV effect. The formulation we developed in this paper can be straightforwardly extended to the halo-shear bispectra such as the halo-shear-shear and halo-halo-shear three-point correlation functions.

Finally, there are several effects that we ignored in this paper and need to be studied. First, we only considered a simple circular-shaped survey geometry to compute the HSV effect. As we showed (see § 3.4), the HSV depends on the survey window function, which in turn determines which large-scale density fluctuation modes, outside the survey region, contribute to the error covariances. For a general-shape survey geometry, the Fourier-transformed survey window function becomes anisotropic, and causes additional apparent correlations between the different spectra, in a two-dimensional way of Fourier space. Hence, numerical simulations would be needed to study this effect. Since the large-scale density fluctuations, which contribute to the HSV effect, are at very large-length scales and therefore in the linear regime, it may be straightforward to take into account the survey window effects, by using our formulation. Secondly, in this paper, we used the flat-sky approach, which is not valid for a very wide solid-angle survey such as LSST. For a full-sky survey as the ultimate case, the HSV effect is caused by the horizon-scale perturbations, outside the survey region, and the fluctuation modes are at wavenumbers comparable to the matter-radiation equality wavenumber k_{eq} . In this case, the HSV effect decreases more rapidly with the survey volume than other covariance terms ($\propto \Omega_s$) as we discussed below Eq. (14). Hence the HSV effect may not be that significant for such an all-sky survey. This is worth studying, and the formulation and method we developed in this paper can be used for such a study.

ACKNOWLEDGMENTS

We would like to thank Gary Bernstein, Elisabeth Krause, Takahiro Nishimichi and Masanori Sato. This work is supported in part by JSPS KAKENHI (Grant Number: 23340061 and 24740171), by JSPS Core-to-Core Program “International Research Network for Dark Energy”, by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan, and by the FIRST program “Subaru Measurements of Images and Redshifts (SuMIRe)”, CSTP, Japan, and by NSF grant AST-0908027 and DOE grant DE-FG02-95ER40893.

APPENDIX A: DERIVATION OF THE BISPECTRUM COVARIANCE

We derive equations of the bispectrum covariance matrix given in terms of the lensing spectra (power spectrum, bispectrum and the higher-order functions), following the method developed in Takada & Bridle (2007).

A1 Discrete Fourier decomposition

The lensing power spectra need to be estimated from the Fourier-transform of the lensing convergence field, $\tilde{\kappa}_{\mathbf{l}}$. The Fourier decomposition has to be done for modes taken from a finite survey region. For such a finite sky measurement, an infinite number of the Fourier modes are not available. Therefore, the Fourier decomposition is by nature discrete, and the fundamental mode is limited by the size of surveyed area, $l_f = 2\pi/\Theta_s$, where the survey area is given by $\Omega_s = \Theta_s^2$ (we assume a square survey geometry for simplicity). For this case, the convergence field can be expanded using the discrete Fourier decomposition as

$$\kappa(\boldsymbol{\theta}) = \frac{1}{\Omega_s} \sum_{\mathbf{l}} \tilde{\kappa}_{\mathbf{l}} e^{i\mathbf{l} \cdot \boldsymbol{\theta}}, \quad (\text{A1})$$

where the summation runs over the combination of integers (n_x, n_y) for $\mathbf{l} = (2\pi/\Theta_s)(n_x, n_y)$. The prefactor $1/\Omega_s$ is our convention, motivated by the fact that the discrete Fourier decomposition has the limit of the continuous Fourier decomposition; $(1/\Omega_s) \sum_{\mathbf{l}} \tilde{\kappa}_{\mathbf{l}} e^{i\mathbf{l} \cdot \boldsymbol{\theta}} \rightarrow \int d^2\mathbf{l}/(2\pi)^2 \tilde{\kappa}_{\mathbf{l}} e^{i\mathbf{l} \cdot \boldsymbol{\theta}}$ for the limit of $\Theta_s \rightarrow \infty$. Thus, in the discrete Fourier decomposition, the lensing field is expanded by the Fourier modes confined within the finite survey region. In other words, this treatment cannot include the modes of scales outside the survey region.

For this reason, the discrete Fourier decomposition formulation cannot yield the HSV effect that arises from the density fluctuations of scales outside the survey region.

In the discrete Fourier decomposition, the orthogonal relation for eigen function $e^{i\mathbf{l} \cdot \boldsymbol{\theta}}$ is modified as

$$\int_{\Omega_s} d^2\boldsymbol{\theta} e^{i(\mathbf{l}-\mathbf{l}') \cdot \boldsymbol{\theta}} = \Omega_s \delta_{\mathbf{l}-\mathbf{l}'}^K, \quad (\text{A2})$$

where the integration range is confined to the survey region, and $\delta_{\mathbf{l}}^K$ is the Kronecker-type delta; $\delta_{\mathbf{l}}^K = 1$ if $\mathbf{l} = \mathbf{0}$, and otherwise $\delta_{\mathbf{l}}^K = 0$. The orthogonal relation above suggests that the Kronecker delta should be replaced with the Dirac delta function for the limit of $\Theta_s \rightarrow \infty$: $\Omega_s \delta_{\mathbf{l}-\mathbf{l}'}^K \rightarrow (2\pi)^2 \delta^D(\mathbf{l} - \mathbf{l}')$.

Hence the definitions of the lensing spectra are modified for a finite-area survey from Eq. (7); e.g.

$$\begin{aligned} \langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \rangle &= \Omega_s \delta_{\mathbf{l}_1+\mathbf{l}_2}^K P(l_1) \\ \langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \tilde{\kappa}_{\mathbf{l}_3} \rangle &= \Omega_s \delta_{\mathbf{l}_1+\mathbf{l}_2+\mathbf{l}_3}^K B(l_1, l_2, l_3), \end{aligned} \quad (\text{A3})$$

and similar expressions for the higher-order functions. These expressions have the limits; $\langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \rangle = \Omega_s \delta_{\mathbf{l}_1+\mathbf{l}_2}^K P(l_1) \rightarrow \langle \tilde{\kappa}_{\mathbf{l}_1} \tilde{\kappa}_{\mathbf{l}_2} \rangle = (2\pi)^2 \delta^D(\mathbf{l}_1 + \mathbf{l}_2) P(l_1)$ and so on for an infinite-area limit ($\Omega_s \rightarrow \infty$).

A2 Bispectrum estimator and the covariance matrix

The bispectrum is given as a function of triangle configurations. We use a parametrization of three side lengths (l_1, l_2, l_3) to specify a triangle configuration. The three parameters are enough, because all the triangles transformed by parallel transformation, rotation, permutation and parity transformations in Fourier space are equivalent to yielding the same bispectrum in an ensemble average sense for a homogeneous and isotropic field as expected for the lensing field. Hence, within the framework of the discrete Fourier decomposition we described above, we can define an estimator of the bispectrum:

$$\hat{B}(l_1, l_2, l_3) = \frac{1}{\Omega_s N_{\text{trip}}(l_1, l_2, l_3)} \sum_{\mathbf{q}_i} \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3), \quad (\text{A4})$$

where we have introduced the abbreviated notation $\mathbf{q}_{123} \equiv \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$ and the summation runs over all the grids of $\mathbf{q}_1, \mathbf{q}_2$ and \mathbf{q}_3 (3 dimensional summation) in Fourier space. The function $\Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3)$ denotes the selection function defined so that $\Delta_{\mathbf{q}_{123}} = 1$ if each vector has the target length, $l_i - \Delta l_i/2 \leq q_i \leq l_i + \Delta l_i/2$ ($i = 1, 2, 3$) as well as the three vectors satisfy the triangle condition, $\mathbf{q}_{123} = \mathbf{0}$; otherwise $\Delta_{\mathbf{q}_{123}} = 0$. The quantity N_{trip} is the number of triplets of grids (vectors) (see Fig. 3) which form the triangle configuration of (l_1, l_2, l_3).

As implied from Eq. (A4) and Fig. 3, we need to average the estimator of $\tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3}$ over all the triangles that have the side lengths of (l_1, l_2, l_3) within the bin widths. For the limits of $l_i \gg l_f$, the number of independent triplets of a target triangle configuration, N_{trip} , can be estimated as

$$\begin{aligned} N_{\text{trip}}(l_1, l_2, l_3) &\equiv \sum_{\mathbf{q}_i} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3), \\ &\approx 2 \times \frac{(2\pi l_1 \Delta l_1) \times (l_2 \Delta \varphi_{12} \times \Delta l_2)}{(2\pi/\Theta_s)^4} = \frac{\Omega_s^2}{4\pi^3} l_1 l_2 \Delta l_1 \Delta l_2 \Delta \varphi_{12}, \end{aligned} \quad (\text{A5})$$

where we meant by ‘‘independent triplets’’ different combinations of triplets that form the target triangle configuration, and Δl_i is the bin width around each multipole bin, φ_{12} is the angle bounded by the two side lengths q_1 and q_2 in Fig. 3, and $\Delta \varphi_{12}$ is the angle extent allowed by the multipole bins as we explain below. First, $2\pi l_1 \Delta l_1 / (2\pi/\Theta_s)^2$ is the number of grids in the annulus, which has the radius of l_1 with width Δl_1 , as in the coefficient in the power spectrum covariance (see Eq. 12). Then, for the first grid of the vector \mathbf{q}_1 , we can find other two vertices (two grids), denoted by the vectors \mathbf{q}_2 and \mathbf{q}_3 in Fig. 3, so that the two vectors have the target lengths, $q_2 = l_2$ and $q_3 = l_3$ within the bin widths. Given the bin widths, the number of grids, which are covered by variations in the two vectors \mathbf{q}_2 and \mathbf{q}_3 , can be estimated as $(l_2 \Delta \varphi_{12} \times \Delta l_2) / (2\pi/\Theta_s)^2$, where $\Delta \varphi_{12}$ is the variation in the angle φ_{12} allowed by the variations in the two vectors \mathbf{q}_1 and \mathbf{q}_2 . Finally, the prefactor 2 in Eq. (A5) accounts for counter-part triplet of grids that are transformed by parity transformation $\varphi_{12} \rightarrow -\varphi_{12}$.

As can be found from Figure 3, the angle φ_{12} is given as

$$\cos \varphi_{12} = \frac{q_1^2 + q_2^2 - q_3^2}{2q_1 q_2}, \quad (\text{A6})$$

where the two vectors have the target lengths within the bin widths: $l_i - \Delta l_i/2 \leq q_i \leq l_i + \Delta l_i/2$ ($i = 1, 2, 3$). If varying the side length q_3 by Δl_3 , we can find the angle variation as

$$\Delta \varphi_{12} \simeq (\sin \varphi_{12})^{-1} \frac{l_3 \Delta l_3}{l_1 l_2} = \frac{2l_3 \Delta l_3}{\sqrt{2l_1^2 l_2^2 + 2l_1^2 l_3^2 + 2l_2^2 l_3^2 - l_1^4 - l_2^4 - l_3^4}}. \quad (\text{A7})$$

Therefore, substituting this equation into Eq. (A5) leads us to find the number of independent triplets that form the triangle configuration of (l_1, l_2, l_3) :

$$N_{\text{trip}}(l_1, l_2, l_3) = \frac{\Omega_s^2}{2\pi^3} \frac{l_1 l_2 l_3 \Delta l_1 \Delta l_2 \Delta l_3}{\sqrt{2l_1^2 l_2^2 + 2l_1^2 l_3^2 + 2l_2^2 l_3^2 - l_1^4 - l_2^4 - l_3^4}}. \quad (\text{A8})$$

The triplet number has a symmetric property under permutation of $l_1 \leftrightarrow l_2$ and so on. As shown in Joachimi et al. (2009), this expression gives a good approximation, better than 1% accuracy, to the full-sky expression that is given by the Wigner-3j symbol, for the limits of $l_i \gg 1$ (i.e. the flat-sky limit).

The ensemble average of the bispectrum estimator (A4) is found to yield the lensing bispectrum:

$$\begin{aligned} \langle \hat{B}(l_1, l_2, l_3) \rangle &= \frac{1}{\Omega_s N_{\text{trip}}(l_1, l_2, l_3)} \sum_{\mathbf{q}_i} \langle \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \rangle \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) \\ &= \frac{1}{\Omega_s N_{\text{trip}}(l_1, l_2, l_3)} \sum_{\mathbf{q}_i} B(l_1, l_2, l_3) \Omega_s \delta_{\mathbf{q}_{123}}^K \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) \\ &\simeq \frac{1}{N_{\text{trip}}} B(l_1, l_2, l_3) \sum_{\mathbf{q}_i} \delta_{\mathbf{q}_{123}}^K \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) \\ &= B(l_1, l_2, l_3), \end{aligned} \quad (\text{A9})$$

where we have assumed on the third line that the lensing bispectrum does not largely change within the multipole bins Δl_i , and on the forth line that the Kronecker delta $\delta_{\mathbf{q}_{123}}^K$ automatically holds together with the selection function $\Delta_{\mathbf{q}_{123}}$; i.e., $\delta_{\mathbf{q}_{123}}^K = 1$ when $\Delta_{\mathbf{q}_{123}} = 1$.

The bispectrum covariance is defined as

$$\text{Cov} [B(l_1, l_2, l_3), B(l'_1, l'_2, l'_3)] \equiv \frac{1}{\Omega_s N_{\text{trip}}} \frac{1}{\Omega_s N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} \langle \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \tilde{\kappa}_{\mathbf{q}'_1} \tilde{\kappa}_{\mathbf{q}'_2} \tilde{\kappa}_{\mathbf{q}'_3} \rangle \Delta_{\mathbf{q}_{123}} \Delta_{\mathbf{q}'_{123}} - B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3), \quad (\text{A10})$$

where we omitted the arguments in N_{trip} and $\Delta_{\mathbf{q}_{123}}$ such as $N'_{\text{trip}} \equiv N_{\text{trip}}(l'_1, l'_2, l'_3)$ for notational simplicity. Thus the bispectrum covariance arises from the 6-point correlation function.

The 6-point correlation function in can be further computed as

$$\begin{aligned} \langle \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \tilde{\kappa}_{\mathbf{q}'_1} \tilde{\kappa}_{\mathbf{q}'_2} \tilde{\kappa}_{\mathbf{q}'_3} \rangle &= \Omega_s^3 P(q_1) P(q_3) P(q'_2) \delta_{\mathbf{q}_{12}}^K \delta_{\mathbf{q}_{31'}}^K \delta_{\mathbf{q}_{2'3'}}^K + 14 \text{ perms.} \\ &+ \Omega_s^2 B(q_1, q_2, q_3) B(q'_1, q'_2, q'_3) \delta_{\mathbf{q}_{123}}^K \delta_{\mathbf{q}_{1'2'3'}}^K + 9 \text{ perms.} \\ &+ \Omega_s^2 P(q_1) T(\mathbf{q}_3, \mathbf{q}'_1, \mathbf{q}'_2, \mathbf{q}'_3) \delta_{\mathbf{q}_{12}}^K \delta_{\mathbf{q}_{31'2'3'}}^K + 14 \text{ perms.} \\ &+ \Omega_s T_6(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}'_1, \mathbf{q}'_2, \mathbf{q}'_3) \delta_{\mathbf{q}_{1231'2'3'}}^K, \end{aligned} \quad (\text{A11})$$

where $\delta_{\mathbf{q}_{1231'}}^K \equiv \delta_{\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}'_1}^K$ and so on. Inserting the 6-point correlation function above into Eq. (A10) yields

$$\begin{aligned} \text{Cov} [B(l_1, l_2, l_3), B(l'_1, l'_2, l'_3)] &= \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) \Delta_{\mathbf{q}'_{123}}(l'_1, l'_2, l'_3) \\ &\times \left\{ \Omega_s P(q_1) P(q_2) P(q_3) [\delta_{\mathbf{q}_{11'}}^K \delta_{\mathbf{q}_{22'}}^K \delta_{\mathbf{q}_{33'}}^K + \delta_{\mathbf{q}_{11'}}^K \delta_{\mathbf{q}_{23'}}^K \delta_{\mathbf{q}_{32'}}^K + \delta_{\mathbf{q}_{12'}}^K \delta_{\mathbf{q}_{21'}}^K \delta_{\mathbf{q}_{33'}}^K + 3 \text{ perms.}] \right. \\ &+ B(q_1, q_2, q'_1) B(q_3, q'_2, q'_3) \delta_{\mathbf{q}_{121'}}^K \delta_{\mathbf{q}_{32'3'}}^K + 8 \text{ perms.} \\ &+ P(q_1) T(\mathbf{q}_2, \mathbf{q}_3, \mathbf{q}'_2, \mathbf{q}'_3) \delta_{\mathbf{q}_{11'}}^K \delta_{\mathbf{q}_{232'3'}}^K + 8 \text{ perms.} \\ &\left. + \frac{1}{\Omega_s} T_6(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}'_1, \mathbf{q}'_2, \mathbf{q}'_3) \delta_{\mathbf{q}_{1231'2'3'}}^K \right\}, \end{aligned} \quad (\text{A12})$$

where we have dropped terms including the Kronecker delta such as $\delta_{\mathbf{q}_{12}}^K$, because these conflict with the triangle conditions $\mathbf{q}_{123} = \mathbf{0}$ or $\mathbf{q}_{1'2'3'} = \mathbf{0}$ (these imply $q_3 = 0$ if $\mathbf{q}_{12} = \mathbf{0}$). As we explained in Fig. 1, the bispectrum covariance has different contributions that arise from the terms proportional to P^3 , $O(BB)$, $O(PT)$ and P_6 , respectively. We below show further simplified expressions of each term in the bispectrum covariance.

A2.1 Gaussian error contribution to the bispectrum covariance

First, we consider the term in Eq. (A12) proportional to the power spectra cubed P^3 . Hence we call this term the Gaussian error contribution, even though the bispectrum itself is a measure of non-Gaussianity in the lensing field. In fact, many previous works only considered this term when studying the lensing bispectrum, e.g. for parameter forecasts, mainly for simplicity. The Gaussian term can be further simplified as

$$\text{COV}_{\text{Gauss}} = \frac{\Omega_s}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} \Delta_{\mathbf{q}_{123}}(l_1, l_2, l_3) \Delta_{\mathbf{q}'_{123}}(l'_1, l'_2, l'_3) P(q_1) P(q_2) P(q_3) [\delta_{\mathbf{q}_{11'}}^K \delta_{\mathbf{q}_{22'}}^K \delta_{\mathbf{q}_{33'}}^K + 5 \text{ perms.}]$$

$$\begin{aligned}
&\simeq \frac{\Omega_s}{N_{\text{trip}} N'_{\text{trip}}} \sum_{\mathbf{q}_i} \Delta \mathbf{q}_{123}(l_1, l_2, l_3) P(q_1) P(q_2) P(q_3) \left[\delta_{l_1 l'_1}^K \delta_{l_2 l'_2}^K \delta_{l_3 l'_3}^K + 5 \text{ perms.} \right] \\
&= \frac{\Omega_s}{N_{\text{trip}}} P(l_1) P(l_2) P(l_3) \left[\delta_{l_1 l'_1}^K \delta_{l_2 l'_2}^K \delta_{l_3 l'_3}^K + 5 \text{ perms.} \right], \tag{A13}
\end{aligned}$$

where we have used the facts that the power spectrum does not largely change within the bin width and that the term including the Kronecker delta (the terms in the square bracket) is *non-vanishing* only if the two triangles of the two bispectra have the *same* shape. The term including the Kronecker delta give 1, 2 or 6 for general, isosceles and equilateral triangles, respectively, which corresponds to the factor Δ in Eq. (28) in Takada & Jain (2004). Thus the Gaussian error covariance term contributes only to the diagonal terms of the bispectrum covariance matrix. Recalling the fact $N_{\text{trip}} \propto 1/\Omega_s^2$ (Eq. A8), this contribution scales with survey area as $\text{Cov}_{\text{Gauss}} \propto 1/\Omega_s$.

A2.2 Non-Gaussian error contributions to the bispectrum covariance

Other terms in Eq. (A12) are non-Gaussian error contributions to the bispectrum covariance that arise from the higher-order functions. In the following, we derive further simplified expressions of each terms.

A2.2.1 Non-Gaussian terms of $O(B^2)$ We consider the terms in Eq. (A12) that are proportional to the bispectra squared, the terms of $O(B^2)$. Here we consider the first term of $O(B^2)$ terms in Eq. (A12) as an example as

$$\begin{aligned}
&\frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} B(q_1, q_2, q'_1) B(q_3, q'_2, q'_3) \delta_{\mathbf{q}_{123}}^K \delta_{\mathbf{q}'_{123}}^K \Delta \mathbf{q}_{123}(l_i) \Delta \mathbf{q}'_{123}(l'_i) \\
&\simeq \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} B(l_1, l_2, l'_1) B(l_3, l'_2, l'_3) \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} \delta_{\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}'_1}^K \delta_{\mathbf{q}_3 + \mathbf{q}'_2 + \mathbf{q}'_3}^K \Delta \mathbf{q}_{123}(l_i) \Delta \mathbf{q}'_{123}(l'_i) \\
&= \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} B(l_1, l_2, l'_1) B(l_3, l'_2, l'_3) \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \Delta \mathbf{q}_{123}(l_i) \delta_{l_3 l'_1}^K \left\{ \sum_{\mathbf{q}'_2, \mathbf{q}'_3} \delta_{\mathbf{q}_3 + \mathbf{q}'_2 + \mathbf{q}'_3}^K \Delta \mathbf{q}_{123}(l'_i) \right\} \\
&= \frac{1}{N_{\text{trip}}} \frac{\delta_{l_3 l'_1}^K}{N'_{\text{trip}}} B(l_1, l_2, l'_1) B(l_3, l'_2, l'_3) \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \Delta \mathbf{q}_{123}(l_i) \left\{ \sum_{\mathbf{q}'_2, \mathbf{q}'_3} \Delta \mathbf{q}_{123}(l'_i) \right\} \\
&\simeq \frac{1}{N_{\text{trip}}} \frac{\delta_{l_3 l'_1}^K}{N'_{\text{trip}}} B(l_1, l_2, l'_1) B(l_3, l'_2, l'_3) \sum_{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3} \Delta \mathbf{q}_{123}(l_i) \left[\frac{2 \Delta l'_2 \Delta l'_3 \Delta \varphi_{32'}}{(2\pi/\Theta_s)^2} \right] \\
&= \delta_{l_3 l'_1}^K B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3) \frac{\Omega_s}{(2\pi)^2} \frac{2 \Delta l'_2 \Delta l'_3 \Delta \varphi_{32'}}{N'_{\text{trip}}(l'_1, l'_2, l'_3)} \\
&= \delta_{l_3 l'_1}^K \frac{2\pi}{\Omega_s l'_1 \Delta l'_1} B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3). \tag{A14}
\end{aligned}$$

In the above calculation, we have used several simplifications using the triangle conditions and the Kronecker deltas. In the second line of the r.h.s., the product of Kronecker deltas, $\delta_{\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}'_1}^K \delta_{\mathbf{q}_3 + \mathbf{q}'_2 + \mathbf{q}'_3}^K$, is non-vanishing only if $l_3 = l'_1$, because of the conditions imposed by the selection functions $\Delta \mathbf{q}_{123}(l_i)$ and $\Delta \mathbf{q}'_{123}(l'_i)$; i.e. $l_1 - \Delta l_1/2 \leq q_i \leq l_1 + \Delta l_1/2$, $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = \mathbf{0}$, and so on. In other words, the term above is non-vanishing only if the two triangle configurations have the same length on their one side, $l_3 = l'_1$ in this case (see Fig. 1). Hence, in the second line, we introduced the Kronecker delta $\delta_{l_3 l'_1}^K$, and also used the Kronecker delta $\delta_{\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}'_1}^K$ to drop the summation over \mathbf{q}'_1 (then we also used the triangle conditions, $\mathbf{q}'_1 = -\mathbf{q}_1 - \mathbf{q}_2 = \mathbf{q}_3$ due to $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 = \mathbf{0}$). In the third line, we dropped one Kronecker delta $\delta_{\mathbf{q}_3 + \mathbf{q}'_2 + \mathbf{q}'_3}^K$ because it is automatically satisfied by the selection function $\Delta \mathbf{q}_{123}(l'_i)$. The curly bracket is intended to mean that the summation $\sum_{\mathbf{q}'_2, \mathbf{q}'_3} \Delta \mathbf{q}_{123}(l'_i)$ is done before another summation $\sum_{\mathbf{q}_i}$. In the forth line, we computed the summation $\sum_{\mathbf{q}'_2, \mathbf{q}'_3} \Delta \mathbf{q}_{123}(l'_i)$, which gives the number of triplets of grids satisfying the selection function. For the limits of $q'_2, q'_3 \gg l_f$, we can use the similar calculations given by Eqs. (A5) and (A8). With the selection function $\Delta \mathbf{q}_{123}(l'_i)$, we can find that, for a given vector \mathbf{q}_3 , the summation gives the number of paired grids $(\mathbf{q}'_2, \mathbf{q}'_3)$ within the bin widths, $l'_2 - \Delta l'_2/2 \leq q'_2 \leq l'_2 + \Delta l'_2/2$ and $l'_3 - \Delta l'_3/2 \leq q'_3 \leq l'_3 + \Delta l'_3/2$. According to the similar calculation (Eq. A7), the number of the paired grids is estimated as $2 \times (l'_2 \Delta l'_2 \Delta \varphi_{32'}) / (2\pi/\Theta_s)^2$, where $\Delta \varphi_{32'}$ is the variation in the angle bounded by the two side lengths l'_2 and l'_3 in the triangle of (l_3, l'_2, l'_3) , due to the variations of l'_2 and l'_3 within the bin widths. In the fifth line, we used $\sum_{\mathbf{q}_i} \Delta \mathbf{q}_{123} = N_{\text{trip}}(l_1, l_2, l_3)$ (see Eq. A5). In this step, using the fact $l_3 = l'_1$ via the Kronecker delta $\delta_{l_3 l'_1}^K$, we also replaced $B(l_1, l_2, l'_1) B(l_3, l'_2, l'_3)$ with $B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3)$, the product of the original bispectra taken in the covariance calculation. In the sixth line, we further simplified $2(l'_2 \Delta l'_2 \Delta \varphi_{32'}) / [N_{\text{trip}}(l'_1, l'_2, l'_3) (2\pi/\Theta_s)^2]$ by using the similar equations to Eqs. (A7) and (A8) and the fact $l_3 = l'_1$ to obtain the coefficient $2\pi/[\Omega_s l'_1 \Delta l'_1]$.

Performing the similar calculations to other terms in Eq. (A12), the terms of $O(B^2)$ in the bispectrum covariance can be reduced to

$$\text{Cov}[B(l_i), B(l'_i)]_{\text{NG}, O(B^2)} = \frac{2\pi}{\Omega_s} B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3) \left[\frac{\delta_{l_1 l'_1}^K}{l_1 \Delta l_1} + \frac{\delta_{l_1 l'_2}^K}{l_1 \Delta l_1} + 7 \text{ perms.} \right]. \quad (\text{A15})$$

These terms scale with survey area as $\text{Cov}_{\text{NG}, O(B^2)} \propto 1/\Omega_s$. The terms contribute to diagonal terms of the bispectrum covariance as well as some off-diagonal terms when the two triangle configurations have the same length within the bin widths at least on their one side (see Fig. 1 for the diagrammatic illustration).

A2.2.2 Non-Gaussian terms of $O(PT)$ Next let's consider the terms proportional to $O(PT)$ in Eq. (A12). Similarly to the calculation in Eq. (A14), the first term of Eq. (A12) can be simplified as

$$\begin{aligned} & \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} P(q_1) T(\mathbf{q}_2, \mathbf{q}_3, \mathbf{q}'_2, \mathbf{q}'_3) \delta_{\mathbf{q}_{11'}}^K \delta_{\mathbf{q}_{232'3'}}^K \Delta \mathbf{q}_{123}(l_i) \Delta \mathbf{q}'_{123}(l'_i) \\ & \simeq \delta_{l_1 l'_1}^K P(l_1) T(l_2, l_3, l'_2, l'_3) \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \Delta \mathbf{q}_{123}(l_1, l_2, l_3) \left\{ \sum_{\mathbf{q}'_2, \mathbf{q}'_3} \Delta_{-\mathbf{q}_1 + \mathbf{q}'_2 + \mathbf{q}'_3} \right\} \\ & \simeq \delta_{l_1 l'_1}^K P(l_1) T(l_2, l_3, l'_2, l'_3) \frac{1}{N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \frac{2 \times (l'_2 \Delta l_2 \Delta \varphi_{12'})}{(2\pi/\Theta_s)^2} \sum_{l_i} \Delta l_{123} \\ & \simeq \delta_{l_1 l'_1}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_2, l'_3). \end{aligned} \quad (\text{A16})$$

In the first line of the r.h.s. of the above equation, we used the fact that the term is non-vanishing only if the two triangle configurations have the same length on their one side, $l_1 = l'_1$ in this case (see Fig. 1). In this step, therefore, we dropped one summation over $\sum \mathbf{q}'_1$ by using the Kronecker delta $\delta_{\mathbf{q}_1 + \mathbf{q}'_1}^K$ and introduced the Kronecker delta $\delta_{l_1 l'_1}^K$. We take out the trispectrum from the summation assuming it does not largely change within the bin widths, because the 4-point configuration is uniquely specified by the two triangle configurations; the 4 outer circumference side lengths are given by l_2, l_3, l'_2 and l'_3 , and the diagonal length is given by l_1 ¹². In the second line, we computed the number of the modes given by the summation $\sum_{\mathbf{q}'_2, \mathbf{q}'_3} \Delta_{-\mathbf{q}_1 + \mathbf{q}'_2 + \mathbf{q}'_3}(l'_1, l'_2, l'_3)$ as we did in the calculation of Eq. (A14). The angle $\varphi_{12'}$ is the angle bounded by the two side lengths l_1 and l'_2 , and $\Delta \varphi_{12'}$ is the variation due to the bin widths. In the third line, we used the fact $\sum_{\mathbf{q}_i} \Delta \mathbf{q}_{123} = N_{\text{trip}}(l_i)$ and further simplified the coefficient $2 \times (l'_2 \Delta l_2 \Delta \varphi_{12'})/N_{\text{trip}}(l'_i)$ as in Eq. (A14).

Hence, the terms of $O(PT)$ in the bispectrum covariance are computed for the limits of $l_i, l'_i \gg l_f$ as

$$\text{Cov}[B(l_i), B(l'_i)]_{\text{NG}, O(PT)} = \delta_{l_1 l'_1}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_2, l'_3) + \delta_{l_1 l'_2}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_1, l'_3) + 7 \text{ perms.} \quad (\text{A17})$$

Again note that all the trispectra in the terms are uniquely specified by the two triangle configurations; the 4 outer-circumference side lengths and the diagonal length (see Fig. 1). The terms of $O(PT)$ scale with survey area as $\text{Cov}_{\text{NG}, O(PT)} \propto 1/\Omega_s$. The terms contribute to the diagonal terms of the bispectrum covariance matrix as well as some off-diagonal terms where the two triangles have the same length at least on their one side.

A2.2.3 Non-Gaussian term of $O(P_6)$ Finally we consider the contribution arising from the connected 6-point correlation function in Eq. (A12):

$$\begin{aligned} \text{Cov}[B(l_i), B(l'_i)]_{\text{NG}, O(P_6)} &= \frac{1}{\Omega_s N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}'_i} P_6(\mathbf{q}_1, \dots, \mathbf{q}'_3) \delta_{\mathbf{q}_{1231'2'3'}}^K \Delta \mathbf{q}_{123}(l_i) \Delta \mathbf{q}'_{123}(l'_i) \\ &\simeq \frac{1}{\Omega_s N_{\text{trip}}} \frac{1}{N'_{\text{trip}}} \frac{(2\pi)^{12}}{\Omega_s^6} \int \prod_{i=1}^3 d^2 \mathbf{q}_i \int \prod_{i=1}^3 d^2 \mathbf{q}'_i P_6(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}'_1, \mathbf{q}'_2, \mathbf{q}'_3) \Delta l_{123}(l_1, l_2, l_3) \Delta l'_{123}(l'_1, l'_2, l'_3) \\ &\simeq \frac{1}{\Omega_s} \int \frac{d\psi}{2\pi} P_6(l_1, l_2, l_3, l'_1, l'_2, l'_3; \psi). \end{aligned} \quad (\text{A18})$$

The connected 6-point correlation function depends on the 6-point configuration in Fourier space. As shown in Fig. 1, the two triangle parts of the connected 6 point configuration are specified by the two triangle configurations of the bispectra. The selection functions $\Delta \mathbf{q}_{123}$ and $\Delta \mathbf{q}'_{123}$ automatically satisfies the condition of the 6-point configuration in Fourier space: $\mathbf{q}_{123} + \mathbf{q}'_{123} = \mathbf{0}$. The unspecified part is only the angle between the two triangles; here we introduced the angle ψ , which is the angle between the vectors \mathbf{q}_1 and \mathbf{q}_2 in Fourier space. Hence, the summations $\sum_{\mathbf{q}_i}$ and $\sum_{\mathbf{q}'_i}$ can be replaced with one-dimensional integration over the angle ψ in Fourier space. The covariance term of $O(P_6)$ scales with survey area as $\text{Cov}_{\text{NG}, O(P_6)} \propto 1/\Omega_s$. This term contributes to both the diagonal and off-diagonal parts of the bispectrum covariance.

¹² The two-dimensional trispectrum is uniquely specified by 5 parameters to characterize the 4-point configuration; e.g., the outer-circumference 4 side lengths plus the diagonal length.

A2.3 Flat-sky formula for the bispectrum covariance

Summing up all the terms of Eqs. (A13), (A15), (A17) and (A18), the covariance matrix is found to be given as

$$\begin{aligned} \text{Cov}[B(l_1, l_2, l_3), B(l'_1, l'_2, l'_3)] &= \frac{\Omega_s}{N_{\text{trip}}(l_1, l_2, l_3)} P(l_1)P(l_2)P(l_3) \left[\delta_{l_1 l'_1}^K \delta_{l_2 l'_2}^K \delta_{l_3 l'_3}^K + \delta_{l_1 l'_1}^K \delta_{l_2 l'_3}^K \delta_{l_3 l'_2}^K + \delta_{l_1 l'_2}^K \delta_{l_2 l'_1}^K \delta_{l_3 l'_3}^K + 3 \text{ perms.} \right] \\ &+ \frac{2\pi}{\Omega_s} B(l_1, l_2, l_3) B(l'_1, l'_2, l'_3) \left[\frac{\delta_{l_1 l'_1}^K}{l_1 \Delta l_1} + \frac{\delta_{l_1 l'_2}^K}{l_1 \Delta l_1} + 7 \text{ perms.} \right] \\ &+ \delta_{l_1 l'_1}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_2, l'_3) + \delta_{l_1 l'_2}^K \frac{2\pi}{\Omega_s l_1 \Delta l_1} P(l_1) T(l_2, l_3, l'_1, l'_3) + 7 \text{ perms.} \\ &+ \frac{1}{\Omega_s} \int \frac{d\psi}{2\pi} P_6(l_1, l_2, l_3, l'_1, l'_2, l'_3; \psi). \end{aligned} \quad (\text{A19})$$

However, we should again note that these terms are derived based on the discrete Fourier decomposition formulation, and therefore we include the Fourier modes confined within the survey region, where the Fourier modes obey periodic boundary conditions. Hence, the non-Gaussian error contributions in the above equation account only for the mode coupling between such Fourier eigenmodes. As we have shown in the main part of this paper, we also need to include the halo sample variance (HSV) contribution to the non-Gaussian errors, which arises from a coupling of the modes confined within the survey region with the mass density fluctuations of scales comparable with or larger than the survey region.

APPENDIX B: CROSS-COVARIANCE BETWEEN LENSING POWER SPECTRUM AND BISPECTRUM

Based on the discrete Fourier decomposition formulation, the power spectrum estimator can be defined in Takada & Bridle (2007) as

$$\hat{P}(l) \equiv \frac{1}{\Omega_s N_{\text{pairs}}(l)} \sum_{\mathbf{q}: q \in l} \tilde{\kappa}_{\mathbf{q}} \tilde{\kappa}_{-\mathbf{q}}, \quad (\text{B1})$$

where the summation is over Fourier modes which have the length of l to within the bin width Δl .

The cross-covariance between the power spectrum $P(l)$ and the bispectrum $B(l_1, l_2, l_3)$ is defined as

$$\text{Cov}[P(l), B(l_1, l_2, l_3)] = \frac{1}{\Omega_s N_l} \frac{1}{\Omega_s N_{\text{trip}}} \sum_{\mathbf{q}: q \in l} \sum_{\mathbf{q}_i} \langle \tilde{\kappa}_{\mathbf{q}} \tilde{\kappa}_{-\mathbf{q}} \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \rangle \Delta \mathbf{q}_{123}(l_i) - P(l) B(l_1, l_2, l_3). \quad (\text{B2})$$

Thus the cross-covariance depends on the 5-point correlation function. The 5-point correlation function in the above equation can be further computed as

$$\begin{aligned} \langle \tilde{\kappa}_{\mathbf{q}} \tilde{\kappa}_{-\mathbf{q}} \tilde{\kappa}_{\mathbf{q}_1} \tilde{\kappa}_{\mathbf{q}_2} \tilde{\kappa}_{\mathbf{q}_3} \rangle &= \Omega_s^2 P(q) B(q_1, q_2, q_3) \delta_{\mathbf{q}_{123}}^K + \Omega_s^2 P(q) B(q, q_2, q_3) \delta_{\mathbf{q}+\mathbf{q}_1}^K \delta_{-\mathbf{q}+\mathbf{q}_2+\mathbf{q}_3}^K + 8 \text{ perms.} \\ &+ \Omega_s P_5(\mathbf{q}, -\mathbf{q}, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \delta_{\mathbf{q}_{123}}^K \end{aligned} \quad (\text{B3})$$

Inserting this equation into Eq.(B2) and using the similar calculation we have used, we can find the flat-sky formula for the cross-covariance as

$$\text{Cov}[P(l), B(l_1, l_2, l_3)] = \delta_{ll_1}^K \frac{4\pi}{\Omega_s l_1 \Delta l_1} P(l) B(l_1, l_2, l_3) + 2 \text{ perms.} + \frac{1}{\Omega_s} \int \frac{d\psi}{2\pi} P_5(l, -l, l_1, l_2, l_3; \psi), \quad (\text{B4})$$

where ψ is the angle between the two vectors \mathbf{l} and \mathbf{l}_1 in Fourier space.

REFERENCES

- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Bergé J., Amara A., Réfrégier A., 2010, ApJ, 712, 992
- Bernardeau F., Mellier Y., van Waerbeke L., 2002, A&A, 389, L28
- Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1
- Cooray A., Hu W., 2001a, ApJ, 554, 56
- Cooray A., Hu W., 2001b, ApJ, 548, 7
- Cooray A., Sheth R., 2002, Physics Report, 372, 1
- de Putter R., Wagner C., Mena O., Verde L., Percival W. J., 2012, JCAP, 4, 19
- Dodelson S., Zhang P., 2005, Phys. Rev. D, 72, 083001
- Hamana T., Mellier Y., 2001, MNRAS, 327, 169
- Hamana T. et al., 2003, ApJ, 597, 98
- Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188

- Hikage C., Takada M., Hamana T., Spergel D., 2011, *MNRAS*, 412, 65
- Hoekstra H., Jain B., 2008, *Annual Review of Nuclear and Particle Science*, 58, 99
- Hu W., Kravtsov A. V., 2003, *ApJ*, 584, 702
- Hui L., 1999, *ApJ*, 519, L9
- Huterer D., Takada M., 2005, *Astroparticle Physics*, 23, 369
- Huterer D., Takada M., Bernstein G., Jain B., 2006, *MNRAS*, 366, 101
- Jain B., Khoury J., 2010, *Annals of Physics*, 325, 1479
- Jain B., Seljak U., 1997, *ApJ*, 484, 560
- Jain B., Seljak U., White S., 2000, *ApJ*, 530, 547
- Jarvis M., Bernstein G., Jain B., 2004, *MNRAS*, 352, 338
- Joachimi B., Shi X., Schneider P., 2009, *A&A*, 508, 1193
- Kilbinger M., Schneider P., 2005, *A&A*, 442, 69
- Knox L., 1995, *Phys. Rev. D*, 52, 4307
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Lee J., Pen U.-L., 2008, *ApJ*, 686, L1
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Limber D. N., 1954, *ApJ*, 119, 655
- Ma C., Fry J. N., 2000, *ApJ*, 543, 503
- Martin S., Schneider P., Simon P., 2012, *A&A*, 540, A9
- Miyazaki S. et al., 2006, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Munshi D., Kitching T., Heavens A., Coles P., 2011, *MNRAS*, 416, 1629
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Neyrinck M. C., Szapudi I., Szalay A. S., 2009, *ApJ*, 698, L90
- Oguri M., Takada M., 2011, *Phys. Rev. D*, 83, 023008
- Peacock J. A., Smith R. E., 2000, *MNRAS*, 318, 1144
- Pires S., Leonard A., Starck J.-L., 2012, *MNRAS*, 423, 983
- Rimes C. D., Hamilton A. J. S., 2005, *MNRAS*, 360, L82
- Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, *ApJ*, 701, 945
- Schneider P., 2006, in *Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro*, Meylan G., Jetzer P., North P., Schneider P., Kochanek C. S., Wambsganss J., eds., pp. 269–451
- Scoccimarro R., 2000, *ApJ*, 544, 597
- Scoccimarro R., Frieman J. A., 1999, *ApJ*, 520, 35
- Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, *ApJ*, 546, 20
- Scoccimarro R., Zaldarriaga M., Hui L., 1999, *ApJ*, 527, 1
- Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., 2006, *Phys. Rev. D*, 74, 023522
- Seljak U., 2000, *MNRAS*, 318, 203
- Semboloni E., Heymans C., van Waerbeke L., Schneider P., 2008, *MNRAS*, 388, 991
- Seo H.-J., Sato M., Dodelson S., Jain B., Takada M., 2011, *ApJ*, 729, L11
- Shaw L. D., Zahn O., Holder G. P., Doré O., 2009, *ApJ*, 702, 368
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Smith R. E. et al., 2003, *MNRAS*, 341, 1311
- Springel V., Frenk C. S., White S. D. M., 2006, *Nature*, 440, 1137
- Takada M., Bridle S., 2007, *New Journal of Physics*, 9, 446
- Takada M., Jain B., 2002, *MNRAS*, 337, 875
- Takada M., Jain B., 2003a, *MNRAS*, 340, 580
- Takada M., Jain B., 2003b, *MNRAS*, 344, 857
- Takada M., Jain B., 2004, *MNRAS*, 348, 897
- Takada M., Jain B., 2009, *MNRAS*, 395, 2065
- Takahashi R. et al., 2011, *ApJ*, 726, 7
- Takahashi R. et al., 2009, *ApJ*, 700, 479
- Tegmark M., Taylor A. N., Heavens A. F., 1997, *ApJ*, 480, 22
- Valageas P., Nishimichi T., 2011a, *A&A*, 527, A87
- Valageas P., Nishimichi T., 2011b, *A&A*, 532, A4
- Valageas P., Sato M., Nishimichi T., 2012, *A&A*, 541, A161
- Van Waerbeke L., Hamana T., Scoccimarro R., Colombi S., Bernardeau F., 2001, *MNRAS*, 322, 918

- White M., 2004, *Astroparticle Physics*, 22, 211
White M., Hu W., 2000, *ApJ*, 537, 1
Zhan H., Knox L., 2004, *ApJ*, 616, L75
Zhang P., Sheth R. K., 2007, *ApJ*, 671, 14
Zhang T.-J., Pen U.-L., Zhang P., Dubinski J., 2003, *ApJ*, 598, 818